

申请上海交通大学博士学位论文

SAR 图像目标/地物精细化解译方法研究

论文作者 赵娟萍

学 号 0140349075

导 师 熊惠霖 教授

专 业 信息与通信工程

答辩日期 2020 年 07 月 09 日

Submitted in total fulfillment of the requirements for the degree of Doctor
in Information and Telecommunication Engineering

Research on SAR Image Object/Land Fine Interpretation Methods

JUANPING ZHAO

Advisor

Prof. HUILIN XIONG

SCHOOL OF ELECTRIC INFORMATION AND ELECTRONIC ENGINEERING

SHANGHAI JIAO TONG UNIVERSITY

SHANGHAI, P.R.CHINA

July 9th, 2020

上海交通大学 学位论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全意识到本声明的法律结果由本人承担。

学位论文作者签名：赵焜萍

日期：2020年7月9日

SAR 图像目标/地物精细化解译方法研究

摘 要

合成孔径雷达 (Synthetic Aperture Radar, SAR) 由于具有全天时全天候对地观测能力、较强的穿透能力, 以及极化信息获取能力, 在军事和民用领域中均得到了广泛的研究与应用。SAR 图像解译是通过研究 SAR 图像中目标和地物的特性与相互作用从图像中获取关键信息的技术。近年来, 随着 SAR 图像数据量的急剧增长和卷积神经网络 (Convolutional Neural Networks, CNNs) 等机器学习方法的大规模应用, 单一的技术或方法已经难以满足 SAR 图像精细化解译的需求, SAR 图像解译方法也呈现出了传统模型与深度网络相结合的发展趋势。

本文基于卷积神经网络和高斯过程回归等机器学习方法, 分别从目标检测、极化方式、探测波段、地物库构建与质量评估方面开展 SAR 图像目标与地物精细化解译研究。主要研究内容包括: 面向 SAR 图像舰船目标检测的精细化解译、多种极化方式下 SAR 图像的物理散射类型提取、多种探测波段下 SAR 图像的物理散射类型提取, 以及 SAR 图像精细化解译测试库构建与质量评估。本文的创新性研究成果总结如下:

(1) 本文首先针对 SAR 图像中靠港/岸等密集排布舰船目标检测困难的问题, 采用改进的多尺度卷积神经网络方法, 同时实现了 SAR 图像中远海海域和近港/岸区域舰船目标的高精度检测。其次, 针对检测框不够精确、密集排布目标漏检严重, 及少量陆地虚警的问题, 本文提出了一种结合级联卷积神经网络和频域脉冲余弦变换的方法用于 SAR 图像舰船目标的精细化提取。该方法通过级联卷积神经网络提取 SAR 图像中舰船目标的粗略候选区, 并在局部自适应范围内采用基于脉冲余弦变换的方法进行目标的精细化提取。在欧空局星载 Sentinel-1 和我国高分三号数据集上的实验结果显示了本文方法的优越性。

(2) 为了通过单/双极化 SAR 图像提取具有物理意义的地物散射类型, 并在此基础上探索多种不同极化方式下 SAR 图像对雷达物理散射特性的可解释性, 本文提出了一种复数域对比约束卷积神经网络方法。本文所研究的物理散射类型主要聚焦在多次散射、体散射和表面散射上, 模型训练真值依赖 Cloude 极化目标分解方法自动获取。复数域对比约束卷积神经网络方法能够在对比约束损失函数的监督训练下, 直接从原始复散射矩阵中学习图像中的物理散射类型。该方法能够实现多种极化方式下 SAR 图像的地物物理散射类型提取, 因而能够从极化方式的角度精细化解译雷达物理散射特性。本

上海交通大学 学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权上海交通大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本学位论文属于

保 密 ☐，在 _____ 年解密后适用本授权书。

不保密 ☒。

(请在以上方框内打√)

学位论文作者签名： 赵娟萍

指导教师签名： 熊飞霖

日 期： 2020 年 7 月 9 日

日 期： 2020 年 7 月 13 日

文方法通过德国宇航中心提供的多景 L 波段高分辨率 F-SAR 影像验证了其有效性和泛化性能。

(3) 针对多种探测波段下 SAR 图像中物理散射类型难以有效提取的难题, 本文通过基于高斯过程回归的方法首先预测极化熵和极化散射角等参数, 最终实现跨雷达探测波段的物理散射类型提取。该方法在高斯假设条件下, 首先通过多种探测波段的 SAR 图像训练基于核函数与平均期望的高斯过程回归模型, 所训练的模型能够较为精确地预测极化熵和极化散射角等参数, 并给出预测参数的置信范围; 进而依据 $H-\alpha$ 极化分割平面从多个探测波段的 SAR 图像中提取到具有物理意义的散射类型。本文方法的有效性与泛化性能在 X 波段 TerraSAR-X 影像、C 波段 Sentinel-1 影像、L 波段 F-SAR 影像上得到了验证。

(4) 针对中低分辨率的 Sentinel-1 影像难以实现地物类别精细化解译的问题, 本文在预定义的两级精细化城区类别标注体系基础上, 采用地理位置信息作为纽带, 通过光学影像辅助标注的方式, 构建了一个面向 SAR 图像地物精细化解译的测试库。该测试库能够提供 33,358 个城区地物样本, 每个样本同时对应了两种不同的极化方式 (VH 和 VV) 和四种不同的数据格式 (包括原始 32 位数据、UINT8 格式数据、辐射校正图像和伪彩色图)。本文通过基于快速压缩距离的数据流形可视化方法和多种图像分类方法验证了测试库的质量。

关键词: SAR 图像, 目标/地物精细化解译, 物理散射类型, 卷积神经网络, 高斯过程回归

RESEARCH ON SAR IMAGE OBJECT/LAND FINE INTERPRETATION METHODS

ABSTRACT

With all-day and all-weather image acquisition capability, relatively strong penetration capability, as well as polarimetry characteristics, Synthetic Aperture Radar (SAR) has found various applications in military and civil community. SAR image interpretation is to acquire key information via the interaction between objects and land covers. Recently, with the rapid increase of SAR images and the wide application of machine learning methods represented by Convolutional Neural Networks (CNNs), a single technology or method is difficult to satisfy the requirement of SAR image fine interpretation, which tends to the combination of traditional models and deep CNNs.

Based on multiple machine learning methods, such as CNNs and Gaussian Process Regression (GPR), this thesis aims to research fine interpretation of earth observations for SAR images from the aspects of object detection, polarimetry, detecting wavebands, as well as dataset compilation and quality evaluation. Main contents of this thesis include: object-oriented fine interpretation based on SAR ship detection, polarizations-oriented fine interpretation based on physical scattering characteristics, waveband-oriented fine interpretation based on physical scattering characteristics, and categories-oriented land cover fine interpretation based on dataset compilation and quality evaluation.

Firstly, this thesis realizes high precision SAR ship detection in open sea and nearshore areas by using a revised multi-scale CNN. In order to further attain more accurate detections, reduce missing detections in densely clustered area and false alarms on land, this thesis presents a method combining cascade CNN and Pulse Cosine Transform (PCT) based visual attention method. Aiming to achieve SAR ship candidates, the former analyzes SAR images in the space domain. On the contrary, the latter analyzes SAR images in the frequency domain in adaptive local regions. The superiority of the proposed method has been demonstrated on European Space Agency (ESA)'s Sentinel-1 dataset and the Chinese Gaofen-3 (GF-3) dataset.

Secondly, in order to extract physical signatures from single- and dual-polarimetric SAR images and further explore potentials on SAR images with multiple polarizations, this thesis

proposes a contrastive-regulated CNN in the complex domain, directly learning physically interpretable deep model from complex Sinclair scattering matrices. The input, output, and parameters are all complex values, and the computations are in the complex domain. The loss function is composed of a basic loss and a contrastive regularization term. This study specifically focuses on multiple scattering, volume scattering, and surface scattering. With the assistance of Cloude's polarimetric decomposition, this thesis generates ground-truth of physical scattering types. Experiments on DLR's F-SAR data demonstrate the effectiveness and generality of this method.

Thirdly, this thesis realizes wavebands-oriented fine interpretation for SAR image scattering characteristics based on GPs. Based on Gaussian hypothesis conditions, this thesis proposes a kernel- and expectation-based Gaussian process regression method to learn some polarimetric parameters, such as polarimetric entropy and polarimetric alpha angle. Then, the physical scattering signatures for earth observations can be obtained with the help of $H - \alpha$ division plane. Gaussian processes have the capability of realizing high quality prediction and their corresponding uncertainty levels. The kernel trick and expectation operations could alleviate geometric distortions and calibration discrepancies between images acquired with different wavebands. Experiments on X-band TerraSAR-X images, C-band Sentinel-1 images, and L-band F-SAR images demonstrate the effectiveness and robustness of this method.

Finally, this thesis contributes a large-scale SAR database with relatively low resolution. The images are acquired by Sentinel-1 in 21 Chinese major cities. Based on a pre-defined hierarchical annotation system, the dataset is divided into ten different categories, mainly including building types. The annotation is guaranteed by the transition from optical annotation in Google Earth to SAR annotation by Sentinel-1 Application Platform (SNAP). The database provides 33,358 image patches with 100×100 pixels, each patch is with VH and VV polarizations and four different data formats (including original 32-bit data, UINT8 data, radiometric-calibration data, and pseudo-color data). The quality of this dataset is evaluated via fast compression distance based manifold visualization method and multiple image classification methods.

KEY WORDS: SAR image, object/land fine interpretation, physical scattering signatures, Convolutional Neural Networks (CNNs), Gaussian Process Regression (GPR)

目 录

插图索引	XI
表格索引	XIII
算法索引	XV
第一章 绪论	1
1.1 课题背景与研究意义	1
1.2 国内外研究现状	3
1.2.1 SAR 图像测试库发展	3
1.2.2 SAR 图像分类现状	5
1.2.3 SAR 图像舰船目标检测	7
1.2.4 极化 SAR 图像散射特性	10
1.3 研究内容与组织结构	12
第二章 SAR 与极化 SAR 基础	15
2.1 引言	15
2.2 SAR 成像基础简介	15
2.3 SAR 图像特性	16
2.3.1 几何畸变	17
2.3.2 相干斑噪声	18
2.4 雷达极化原理与特性	19
2.5 极化 SAR 图像基本表征	21
2.5.1 极化散射矩阵及其向量化表征	21
2.5.2 相干矩阵与协方差矩阵	22
2.6 本章小结	23
第三章 面向 SAR 图像舰船目标检测的精细化解译	25
3.1 引言	25
3.2 基于卷积神经网络的目标检测方法概述	26
3.2.1 基于区域候选框预测的双网检测算法	26

3.2.2	基于检测框回归的单网检测算法	27
3.3	基于多尺度卷积神经网络的 SAR 图像舰船目标检测	28
3.3.1	研究动机	28
3.3.2	方法详介	29
3.3.3	实验结果与分析	33
3.4	结合级联网络与脉冲余弦变换的舰船目标精细化提取	40
3.4.1	研究动机	40
3.4.2	方法详介	43
3.4.3	实验结果与分析	51
3.5	本章小结	55
第四章	多种极化方式下 SAR 图像的物理散射类型提取	57
4.1	引言	57
4.2	Cloude 极化目标分解	58
4.3	基于复数域卷积神经网络的物理散射类型提取	60
4.3.1	网络结构	62
4.3.2	网络优化的损失函数	63
4.3.3	物理散射类型提取	66
4.4	实验结果与分析	66
4.4.1	实验数据与网络训练真值	66
4.4.2	实验设置及性能评价	70
4.4.3	多种极化方式下 SAR 图像的物理散射类型提取	71
4.4.4	验证模型的泛化性能	73
4.5	本章小结	76
第五章	多种探测波段下 SAR 图像的物理散射类型提取	79
5.1	引言	79
5.2	高斯过程基本原理及其在遥感图像中的应用	80
5.2.1	高斯过程基本原理	80
5.2.2	高斯过程在遥感图像中的应用	81
5.3	基于高斯过程回归的物理散射类型提取	82
5.3.1	数据准备	82
5.3.2	方法详介	83
5.4	实验分析	88

5.4.1	实验设置	88
5.4.2	基于 X 波段 TerraSAR 数据的物理散射类型提取	89
5.4.3	基于 C 波段 Sentinel-1 数据的物理散射类型提取	91
5.4.4	验证模型的泛化性能	97
5.5	本章小结	97
第六章	SAR 图像精细化解译测试库构建与质量评估	101
6.1	引言	101
6.2	OpenSARUrban 测试库构建方法	102
6.2.1	数据收集与预处理	102
6.2.2	精细化类别体系定义	103
6.2.3	测试库标注方法	105
6.2.4	测试库组织形式	109
6.3	OpenSARUrban 测试库特性分析与规范化管理	110
6.3.1	测试库特性分析	110
6.3.2	测试库规范化管理	113
6.4	测试库质量评估：数据库流形可视化	113
6.5	测试库质量评估：图像分类	114
6.5.1	验证算法与实验设置	115
6.5.2	测试库分类整体性能评价	116
6.6	本章小结	119
全文总结		121
参考文献		123
致 谢		141
攻读学位期间发表的学术论文		143
攻读学位期间参与的项目		147
攻读学位期间参加的学科竞赛及获奖情况		149

插图索引

1-1 传统 SAR 图像分类流程图	5
1-2 传统 SAR 图像舰船检测方法流程图	8
1-3 CFAR 方法的滑动窗口示意图	8
1-4 密集排布目标对 CFAR 窗口的影响	9
1-5 论文研究内容与组织架构	13
2-1 SAR 成像几何关系	16
2-2 合成孔径原理	17
2-3 SAR 图像几何畸变现象	18
2-4 光学图像与 SAR 图像对比	19
2-5 极化 SAR 图像后向散射特性	21
3-1 改进的多尺度卷积神经网络的网络结构图	29
3-2 检测网络候选框提取机制	31
3-3 舰船目标候选区域深度特征与上下文特征融合	32
3-4 宽幅 SAR 图像裁剪方式	35
3-5 在 Sentinel-1 实测 SAR 图像上的 PR 曲线	36
3-6 在 Sentinel-1 影像上新加坡马六甲海峡区域的可视化检测结果	37
3-7 在 Sentinel-1 影像上远海海域的 SAR 图像舰船目标检测结果	38
3-8 在高分三号实测数据上的 PR 曲线	38
3-9 高分三号实测数据上的可视化舰船目标检测结果	40
3-10 在远海海域高分三号实测数据上的 SAR 图像舰船目标检测结果	41
3-11 改进多尺度卷积神经网络方法检测结果中存在的问题分析	41
3-12 结合级联网络与脉冲余弦变换方法用于 SAR 图像舰船目标检测的整体思路	43
3-13 级联卷积神经网络的网络结构图	44
3-14 级联网络模型提取粗略舰船目标候选框	45
3-15 基于脉冲余弦变换的目标精细化提取流程图	48
3-16 局部自适应候选区域设计方法	48

3-17 结合级联网络与脉冲余弦变换的方法在 Sentinel-1 实测数据上的检测效果图	52
3-18 鬼影对舰船目标检测带来的影响	53
3-19 PCT 模型处理的中间结果	54
3-20 相干斑噪声抑制对检测结果的影响	55
4-1 本文所研究的物理散射类型	58
4-2 Cloude 的 $H - \alpha$ 极化分割平面	61
4-3 复数域对比约束卷积神经网络结构图	62
4-4 实验数据预处理与真值	68
4-5 $H - \alpha$ 极化参数分布情况	69
4-6 损失函数项加权系数的选取	71
4-7 从各种极化方式 SAR 图像中提取物理散射类型的准确率比较	72
4-8 单极化 SAR 图像中提取物理散射类型的可视化结果	73
4-9 双极化 SAR 图像中提取物理散射类型的可视化结果	74
4-10 验证模型泛化性能所用的数据及其散射类型真值	74
4-11 从多种极化方式新的 SAR 图像中提取物理散射类型的准确率	75
4-12 单极化 SAR 图像中提取物理散射类型的可视化结果	75
4-13 双极化 SAR 图像中提取物理散射类型的可视化结果	76
5-1 基于高斯过程回归的物理散射类型提取	84
5-2 多种探测波段 SAR 数据覆盖范围	89
5-3 X 波段 TerraSAR 影像极化参数训练过程	90
5-4 通过对数变换的 TerraSAR-X 图像块可视化	90
5-5 X 波段 TerraSAR 数据上预测的极化熵均值和相应的置信水平	92
5-6 X 波段 TerraSAR 数据上模型预测的极化散射角均值及其对应的置信水平	93
5-7 在 TerraSAR-X 影像上的物理散射类型提取结果	94
5-8 基于 C 波段 Sentinel-1 影像的极化参数训练过程	94
5-9 通过对数变换可视化后的 C 波段 Sentinel-1 测试图像	95
5-10 基于 C 波段 Sentinel-1 影像预测的极化熵均值和对应的置信水平	95
5-11 基于 C 波段 Sentinel-1 影像预测的极化散射角均值和对应的置信水平	96
5-12 在 C 波段 Sentinel-1 影像上的物理散射类型识别结果	96
5-13 C 波段宽幅 Sentinel-1 影像的可视化	98
5-14 在 Sentinel-1 影像上的物理散射类型识别结果	99

6-1	OpenSARUrban 测试库数据源分布情况	104
6-2	城区两级类别标注体系	105
6-3	测试库中光学与 SAR 图像样本切片样例对比	106
6-4	OpenSARUrban 数据库构建流程图	108
6-5	重叠面积计算方法	109
6-6	OpenSARUrban 数据库的组织形式	110
6-7	OpenSARUrban 测试库 UINT8 格式切片	111
6-8	测试库影像地缘分布	112
6-9	测试库数据流形可视化方法	114
6-10	OpenSARUrban 测试库流形可视化结果	115
6-11	对整个测试库采用不同评价方法所得的总体准确率	117
6-12	基于 VGG19 网络的不同极化方式测试库分类混淆矩阵	118
6-13	基于 LBP 特征的不同极化方式测试库分类混淆矩阵	118

表格索引

2-1 雷达电磁波波段比较	20
3-1 ESPN 网络中三个预测分支的参数设置	30
3-2 改进的多尺度卷积神经网络结构、参数数量和乘加计算量消耗	33
3-3 在 Sentinel-1 实测数据上的量化结果	36
3-4 在高分三号实测数据上的量化检测结果	39
3-5 结合级联网络与脉冲余弦变换方法的量化检测结果	51
4-1 物理散射特性的真值分配方式	61
4-2 德国 Kaufbeuren 地区 F-SAR 影像基本参数信息	68
5-1 实验数据的基本成像参数	83
6-1 OpenSARUrban 测试库数据来源	103

算法索引

3-1 基于 PCT 变换的视觉关注模型用于 SAR 图像舰船目标鉴别	50
4-1 从单/双极化 SAR 图像中提取具有物理意义的散射类型	67

第一章 绪论

1.1 课题背景与研究意义

遥感技术是 20 世纪 60 年代发展起来的一种远距离新型目标探测技术,它具有非接触、远距离探测、探测面积大,以及获取数据能力强等优点^[1-4]。作为遥感技术中的典型代表,合成孔径雷达 (Synthetic Aperture Radar, SAR) 可以实现稳定的主动式侧视相干成像过程^[5-8]。它通过远距离探测装置雷达天线主动发射并接收多种极化方式和电磁波波段的信号,从而获取地球表面相关状态散射信息数据,进而通过一定的成像技术获取雷达图像。合成孔径技术以雷达天线与地球观测之间的相对运动为基础,解决了为保证更高分辨率而带来的天线孔径过大的问题^[9]。与红外^[10]、光学^[11] 等图像相比, SAR 图像具有其独有的特点。这些特点可以总结为: 1) 具有全天时全天候、成像周期短、连续监测的特点,是在天气情况复杂多变、普通地区夜间时间较长(如极地地区极夜)情况下获取陆地或者海洋信息的重要手段; 2) SAR 具备穿透性强的特点,不受云雨雾等天气条件影响,能够穿透地表、叶簇等地球表面结构; 3) SAR 具备多极化、多频段、多视角获取雷达观测数据的能力。由于其微波波段侧视成像、相位相干处理等特性, SAR 图像与光学图像的成像机理完全不同,不符合人眼视觉系统 (Human Visual System, HVS) 的关注机制。很多情况下,即使是判读专家也很难直观解读 SAR 图像。尽管通过肉眼直观理解 SAR 图像的难度非常大,但是 SAR 图像能够反映地物目标丰富的信息,例如目标结构、几何形状、材质特性等,这些信息都能够通过地球表面物体的电磁后向散射来表征^[12]。

SAR 图像解译是一种通过研究 SAR 图像中地物目标的特性与相互作用,进而从图像中获取关键信息的技术^[13-16]。SAR 图像在军事和民用领域都有着广泛的应用,因而受到了越来越多的关注。在军事领域中, SAR 图像发挥着不可替代的重要作用,能够作为一种用于港口等军事侦查、战场态势评估、辅助决策军事打击事件、发现并反击目标干扰等。在民用领域, SAR 已经广泛应用于森林监测、自动驾驶、城市监管等方面。此外, SAR 图像在海洋遥感、地理信息获取等领域中也都发挥着不可替代的作用。SAR 图像解译是在轨雷达卫星实现成功应用后期的决定性一步,有时候甚至能够在军事打击等事件中起到决策性作用。SAR 图像解译,尤其是针对特定问题的精化解译,需要广泛的专家知识和多种技术相融合,单一的方法或技术很难取得满意的解译效果。

早期的 SAR 系统采用固定的工作状态,工作频段和极化方式都比较单一,散射回波信号对地球表面的刻画与表征性能极为有限。随着雷达极化成像技术和高分辨率成像

技术的发展,雷达图像也向着多频段、多种极化方式等多样化方向发展。具有全极化观测能力的极化 SAR(Polarimetric SAR, PolSAR) 图像作为一种更加先进的雷达观测手段,因其更强的复数据目标信息获取能力,更是获得了各国研究者的广泛关注。与早期的单一极化雷达图像相比,极化 SAR 图像采用多个电磁波频段和极化方式下的复数电磁波散射数据获取更加丰富的地物与目标特性。极化 SAR 图像中不仅包括了地球观测散射回波的幅度、相位、频率,也包含了由电磁波电场矢量变化带来的极化信息。蕴含在极化 SAR 图像中的丰富的极化信息更是难以被肉眼所观测或感知到,这也严重制约了极化 SAR 图像解译应用与研究的发展。

自 1978 年美国发射的第一颗星载 SAR 系统 SEASAT 验证了星载 SAR 能够对舰船目标及其尾迹成像以来,世界各国先后研制并发射了一系列星载 SAR 系统。进入到 21 世纪以来,由于新技术、新体制、新需求等因素的发展与推动,SAR 技术更是迎来了雷达对地观测的黄金发展阶段。世界各国和机构先后发射了一系列星载 SAR 卫星^[17-23]。近期,还涌现出了其他一些 SAR 卫星数据,如加拿大 RADARSAT¹和我国高分三号 (GF3) 卫星²。而且预计未来十年内还将至少有 10 颗星载 SAR 卫星即将发射。随着 SAR 系统的持续发展,各种成像条件、成像模式下的 SAR 数据以爆发式的速度不断丰富,这在很大程度上促进了 SAR 应用需求的持续更新。呈急剧增长趋势的 SAR 数据一方面为 SAR 图像更加精细化的理解提供了丰富的数据基础与支撑;另一方面,也为大数据背景下的 SAR 图像智能化精细化解译提出了更高的要求。如何从海量 SAR 图像中挑选出有价值的信息,从其中提取出有用的信息且能够有效利用这些信息,同时借助先进的技术手段深入研究目标的物理特性、材料特性、几何形状等特征,也对 SAR 图像精细化解译技术提出了更高的要求与挑战。在极化 SAR 发展过程中,美国喷气推进实验室 (Jet Propulsion Laboratory, JPL) 于 20 世纪 80 年代率先研制出了 L 波段、分辨率为 10 米的机载极化 SAR 系统。进入 20 世纪 90 年代之后,德国、加拿大等国家相继对极化 SAR 系统开展深入研究。其中,机载的极化 SAR 系统主要有美国的 AIRSAR/TOPSAR、德国的 E-SAR 与 F-SAR、加拿大的 C/X-SAR、日本的 PI-SAR、丹麦的 EMISAR 等。星载极化 SAR 系统主要有美国的 SIR-C/X-SAR、日本的 ALOS/PALSAR、德国的 TerraSAR-X、加拿大的 RADARSAT-2、我国的 GF3 等。我国在合成孔径雷达技术方面的研究起步较晚,相应的软硬件设施也比较落后。中科院电子学研究所于 1979 年成功研制了第一台机载 SAR 系统。电子科技集团下属第三十八研究所于 2004 年成功研制了首台机载双极化 SAR 系统。我国的第一颗星载 SAR 系统是由中科院电子学研究所于 2006 年研制成功并投入使用的。在 SAR 图像解译研究方面,国内的国防科技大学、复旦大学、上海

¹<https://earth.esa.int/web/eoportal/satellite-missions/r/risat-1>

²http://space.skyrocket.de/doc_sdat/kompsat-5.htm

交通大学、电子科技大学、西安电子科技大学、北京航空航天大学、中科院等高等学府和研究机构也已经开展了大量科技探索和应用导向的研究工作。

自 2012 年以来, 卷积神经网络 (Convolutional Neural Networks, CNNs)^[24, 25] 等深度学习 (Deep Learning, DL) 技术的蓬勃发展, 为自然图像处理带来了突破性的技术革新与应用革新。大数据时代的智能化解译方法, 如计算机视觉理论方法与应用研究, 目前均已取得巨大成功, 相关的理论和工程实践成果不断涌现, 并应用于各个领域, 如人脸检测^[26, 27]、交通系统监测^[28, 29] 等。与此同时, 硬件设备和计算平台性能大幅提升, 海量数据处理和分布式存储技术日趋成熟, 研究人员也开始考虑将新兴技术应用到 SAR 解译中, 从而实现大数据背景下 SAR 图像的智能化解译。面向 SAR 图像解译的深度学习研究方法研究也不断涌现, 该方面内容仍然是当前 SAR 图像解译领域中的重要研究方向之一。这些先进的成像技术和卷积神经网络等深度学习技术的应用与研究都为 SAR 图像精细化解译提供了基本的研究条件。

本文针对 SAR 图像解译当中的特定问题, 基于多种卷积神经网络方法和高斯过程回归等传统机器学习方法开展 SAR 图像中地物与目标的精细化解译研究, 并将研究成果应用于 SAR 图像的定性与定量分析, 具有一定的理论意义和实用价值。

1.2 国内外研究现状

1.2.1 SAR 图像测试库发展

近年来, 随着国家战略发展对卫星任务的需求以及多种多样的传感器技术的发展, 遥感对地观测数据以每天几个太字节 (Trillionbyte, TB) 的速度不断增长^[30]。从数据的高速发展、数据多样性的不断丰富以及数据量的飞速增长上看, SAR 图像解译领域也已经跨入“大数据时代”, 这些海量数据为 SAR 图像解译提供了丰富的数据来源。数据收集能力的增强也使人们开始思考是不是也应该像计算机视觉领域那样建立一个数据样本库, 一方面对收集来的数据进行有序的分类管理; 另一方面, 对 SAR 数据解译算法提供测试与评估的数据基础。在计算机视觉领域, 针对不同的应用建立标准的图像、语音、视频等数据库对相关算法进行测试和评估的做法已经非常普遍, 如 PASCAL 图像识别库¹、FERET 人脸数据库²、MIT 人脸数据库³、卡耐基梅隆大学的 PIE 人脸数据库⁴、FVC 国际指纹识别竞赛库⁵、牛津大学手识别图像库⁶、伊利诺伊州立大学纹理图像数据

¹<https://pjreddie.com/projects/pascal-voc-dataset-mirror/>

²<https://www.nist.gov/itl/products-and-services/color-feret-database>

³<http://cbcl.mit.edu/software-datasets/FaceData2.html>

⁴http://www.ri.cmu.edu/projects/project_418.html

⁵<https://biolab.csr.unibo.it/FVCOnGoing/UI>

⁶<http://www.robots.ox.ac.uk/~vgg/data/hands/>

库¹，以及瑞典皇家理工学院的肢体行为视频数据库²等。

在光学遥感图像解译领域中，学者们已经构建了多个测试库。早在 2010 年，Yang 等人就提供了 UC Merced 土地利用测试库^[31]，该测试库共包含 21 类场景图像，每类提供了 100 张 256×256 像素大小的图片。后来，武汉大学夏桂松教授所在团队相继构建了 AID 数据集^[32] 和 AID++ 数据集^[33]。这两个测试库都是主要提供无人机影像场景切片。其中，AID 数据集提供了 30 类场景切片，共 10,000 个样本切片；AID++ 数据集共包含 400,000 个航空影像样本切片。同年，该团队也贡献了面向光学遥感图像目标检测的 DOTA 数据集^[34]。光学图像由于其直观性、易判读性等天然因素，相应的测试库构建比较容易。与之相反，SAR 图像测试库的构建就显得异常困难。其原因总结为以下几点：1) SAR 的复杂成像系统导致图像中伴随着严重的相干斑噪声；2) SAR 的主动成像特性导致图像中尤其是城市场景中不可避免地存在着叠掩、阴影等几何畸变现象；3) SAR 成像的昂贵代价导致图像的获取，尤其是大规模图像的获取变得异常困难，已标注的测试库也甚少公开。尽管如此，SAR 测试库的研究仍然取得了初步进展。美国航空航天中心（National Aeronautics and Space Administration, NASA）于 1996 年率先公布了一个运动与静止目标数据集（Moving and Stationary Target, MSTAR）^[35]，该数据集包含了不同成像视角下的 10 类军用车辆目标。MSTAR 数据集被广泛应用于 SAR 图像自动目标识别（Automatic Target Recognition, ATR）系统的研究中^[36-38]。由 DAPAR/AFRL 组织管理，全美多个研究团队积极参与的协作研究项目 WinSAR（Western North America InSAR Consortium）³是一个由美国数十所大学与研究机构组建的，利用雷达遥感进行对地观测协作研究的项目。该项目不仅为其成员单位提供美国西部 SAR 数据的收集、整理和共享服务，而且为其成员单位提供雷达遥感软件工具的研发和共享服务。德国的 EADS 公司也构建了一定规模的样本库^[39]：DOSAR 库包含 X 波段 SAR 图像 26 幅，含 1600 个标注的真实目标，地面覆盖范围 135km^2 ；红外图像库含 521 张热红外图像，从不同视角对机场进行拍摄，含 1300 个真实目标，地面覆盖范围 40km^2 。德国宇航中心（German Aerospace Center, DLR）在测试库构建与应用方面也开展了大量研究。Mihai Datcu 团队于 2013 年基于上百景 TerraSAR-X 数据标注了一个大规模的高分辨率 SAR 图像测试库，该测试库共提供了 850 类土地利用类型，覆盖了包括非洲、美洲、中东、欧洲、亚洲等全球 100 多个城市场景^[30]。Mihai Datcu 教授所在团队又于 2018 年标注了一个高分辨率 SAR 图像变化检测测试库，该测试库仍然是基于高分辨率 TerraSAR-X 影像开发^[40]。同年，上海交通大学郁文贤教授课题组在 OpenSAR 平台上公开了其标注的 11,346 个 SAR

¹http://www-cvr.ai.uiuc.edu/ponce_grp/data/

²<http://www.nada.kth.se/cvap/actions/>

³<http://winsar.unavco.org/>

舰船样本切片^[41]，该数据集主要用于 Sentinel-1 影像中的舰船目标解译研究。德国宇航中心朱晓香教授团队构建了 SEN1-2 测试库用于 SAR/光学图像的多模态融合深度学习算法开发^[42]。

1.2.2 SAR 图像分类现状

传统的 SAR 图像分类流程主要涉及四个步骤：图像预处理、特征提取、特征选择与特征融合、分类器选择（如图 1-1 所示）。在 SAR 图像分类过程中，对图像进行预处理是为了后续提取到的不同类别的特征区分性更强，更加突出目标的作用，最常用的图像预处理方法是图像的相干斑滤波。截至目前，研究者们最常用的 SAR 图像相干斑抑制或滤除方法可以分为两大类：一是在成像过程中通过采用多视处理的方式抑制 SAR 图像相干斑噪声；二是对成像所得到的结果进行噪声抑制处理。后者可以根据两种不同的原理通过图像滤波的方式抑制 SAR 图像中的乘性相干斑噪声。第一种途径是通过抑制噪声斑点来改善图像的外观。第二种途径是同时关注其视觉特征和反射系数的估计问题，其中，反射系数正是描述目标物理特性的重要参数。在特征提取问题上，传统方法中一般是通过人工设计特性并通过一定的图像变换操作来实现的，这种人工提取的特征一般对图像描述和表征不够全面和充分。虽然人工提取的特征在一定程度上能够反映图像纹理、性质等，但是也不可避免地会损失图像本身的一些信息；而且，传统 SAR 图像分类流程中的四个步骤是分阶段进行的，这几个步骤之间并不能一致地优化某个整体的目标函数，使得最后优化的参数很难达到最优。

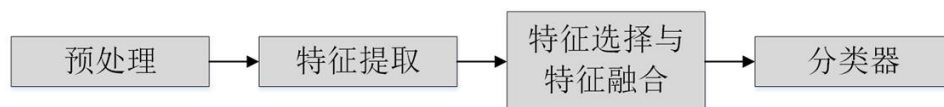


图 1-1 传统 SAR 图像分类流程图

Figure 1-1 Workflow of traditional SAR image classification algorithms

图像特征提取是在原始的“像素级”的图像表示上进行计算建模而得到的向量表达形式，主要包括纹理、形状、空间关系等。特征提取在 SAR 图像分类任务中扮演着非常重要的角色，对图像分类效果起着举足轻重的作用。按照提取范围的不同，图像特征可以分为全局特征和局部特征。全局特征一般通过多维向量的形式对图像中的关键信息进行描述，典型的全局特征包括直方图、纹理能量、轮廓矩等，形式简洁，计算效率较高。但是，由于无法区分前景目标和背景信息，全局特征在描述图像细节信息上的功能比较受限，有时甚至还会丢失图像的空间信息，从而严重影响图像模式分析与分类决策

的准确性。当存在严重的图像形变现象时,全局统计特征差异尤其明显,分类准确性也会呈现明显的下降趋势。因此,全局特征表示方法一般只适用于简单的图像或背景较为单一的图像。与全局特征相比,局部特征的表示方法利用目标局部区域的信息构造特征量。局部信息是指一些在各种成像条件下能够稳定、重复出现,并且具有良好的可区分性的一些点或小范围区域,也称为感兴趣点 (Interesting Point) 或感兴趣区域 (Region of Interest, RoI)。这里所说的局部特征也可以被称作局部不变特征。良好的局部不变特征对复杂背景干扰和部分遮挡条件下的图像处理性能更好,所提取的特征具有可重复性 (Repeatability)、可区分性 (Distinctiveness)、稳健性 (Robustness) 等诸多特点。另外,按照提取方法或途径不同,图像特征也可以分为人工设计特征和机器学习特征。大多数经典的局部不变特征都是采用人工设计与提取的方式获取,需要较多的专家知识。随着图像数据量的增加,以卷积神经网络为代表的机器学习方法大量应用到图像的特征提取中。这些方法大多都是以数据应用为驱动的方法,通过计算机自动学习并挖掘图像中的关键信息。

基于卷积神经网络的 SAR 图像分类是一个端到端的分类学习过程,其本身就是一套系统或者一个架构,而不是一个单一的算法,在这个框架体系中,特征提取和分类是彼此相连、相互作用、相互影响的,遵循预测类别和标准类别之间的差异不断缩小直至收敛的原则,通过损失函数最小化与反向传播算法来不断调整参数,使分类任务朝着预期的方向迭代更新。与传统的分类方法相比,基于卷积神经网络的图像分类具有以下三个优点: 1) 深度结构的卷积神经网络能够通过端到端的训练方式有效表征复杂的目标函数,网络中的所有参数在一个统一的一体化框架中进行优化; 2) 卷积神经网络运用数据驱动的学习方式,直接从海量数据中学习隐含表示,对数据的表征能力更强,因而能够取得更好的性能; 3) 卷积神经网络通过模拟大脑皮层分层处理数据,网络的每层都能提取输入数据不同程度的特征,从而能够建立从底层信号到高层语义特征的映射。在 SAR 图像解译领域,学者们对深度卷积网络应用于 SAR 图像目标分类展开了应用。在 MSTAR 数据集和地物目标分类数据集^[43]上的测试结果显示卷积神经网络能够显著提高 SAR 图像目标分类的准确率,表明卷积神经网络对 SAR 图像同样具有很大的应用潜力。为了解决 MSTAR 数据集规模小和避免过拟合问题,文献 [44, 45] 基于自编码方法^[46]进行卷积网络预训练,文献 [47] 对 MSTAR 原始数据进行变换来增加样本数量,复旦大学金亚秋教授课题组提出了一种全卷积神经网络对 MSTAR 车辆目标进行识别^[36],识别率达到了 99% 以上,取得了目前在 MSTAR 数据集上最好的效果。鉴于复数数据类型在 SAR 图像解译当中的重要作用,Zhang 等提出了一种基于复数据的卷积神经网络^[48],用于对全极化 SAR 图像进行分类,实验结果表明,与实数卷积神经网络相比,复数卷积神经网络方法能够明显提高 SAR 图像分类的准确率。

1.2.3 SAR 图像舰船目标检测

SAR 图像舰船目标检测是从大尺度 SAR 图像中快速准确定位舰船目标位置并区分目标与背景的过程。传统 SAR 舰船检测算法流程图如图 1-2 所示。这类算法流程包括四个关键步骤：图像预处理、海陆分割、舰船检测和舰船目标鉴别。通常，SAR 图像预处理包括图像几何与辐射校正、相干斑噪声抑制等；为了减少不必要的算法计算量和陆地虚警，常用的做法是在检测之前首先对影像进行海陆分割，从而将舰船检测算法集中到感兴趣区域；舰船检测是从感兴趣区域中搜索潜在的疑似舰船目标；舰船目标鉴别是通过典型特征提取与分类的途径剔除虚警。在传统 SAR 图像舰船检测方法中，最经典与常用的一类算法是基于恒虚警率 (Constant False Alarm Rate, CFAR) 的算法。通常，CFAR 方法利用背景区域内的像素根据统计模型自适应地确定阈值。根据恒虚警原则得到判定门限阈值，高于门限阈值的像素点被视为目标像素，否则被视为背景像素，具有自适应阈值与恒虚警率的特性。通过不断滑动窗口，改变背景区域，自适应地求解对应的阈值，最终完成对整幅 SAR 图像的检测。CFAR 目标提取方法中的滑动窗口一般可分为三个部分：目标窗、保护窗、以及背景窗（如图 1-3 所示）。通常，目标窗大小设置为需要检测的舰船目标的尺寸，保护窗大小设置为图像中可能存在的最大舰船目标的尺寸，背景窗应足够大使得背景中的像素数目足够多，确保参数估计的有效性，同时又要避免太大以防止背景窗中包含周围的舰船目标，使得参数估计不准确。但是，在繁忙的港口或者水道区域，大量舰船目标呈现出密集排布的状态，周围的舰船目标会出现在 CFAR 窗口的背景中，并且会干扰背景建模与参数估计。这也是困扰 CFAR 检测算法的一个经典问题。CFAR 背景窗应该足够大，使得背景窗中包含的像素数目足够多，以确保参数估计的有效性；但在舰船目标密集的区域，如果背景窗太大会导致背景中包含周围的舰船目标，将其作为背景来处理，这样会使得参数估计不准确，从而造成目标漏检，如图 1-4 所示。在远海区域，呈现高海况状态的背景区域通常会呈现较亮的状态；同时，由于高海况下的海洋洋流复杂，会使得 SAR 图像整体分布并不均匀，信噪比较低。此外，在中等分辨率和像元大小的情况下，舰船目标通常在图像上表现为很小的区域。这些情况都会增加 CFAR 检测算法在实际应用中的困难。学者们针对不同的应用问题设计了多种衍生算法，包括 CA-CFAR^[49](Cell-average CFAR)，GO-CFAR^[50](Greatest of CFAR)，OS-CFAR^[51](Order-statistic CFAR) 和 VI-CFAR^[52](Variability Index CFAR)。在传统 SAR 图像舰船检测算法中，另一类较为有效的方法是基于广义似然比 (Generalized-Likelihood Ratio) 的方法^[53]。这些多阶段的检测方法算法简单易于实现，但是存在以下缺陷：1) 多阶段处理步骤繁琐，智能化程度不高；2) 参数优化不是在一个框架下进行，使得参数很难达到最优；3) 这些方法严重依赖传统的特征提取方法，而人工特征设计的算法很难全面刻画图像信息。

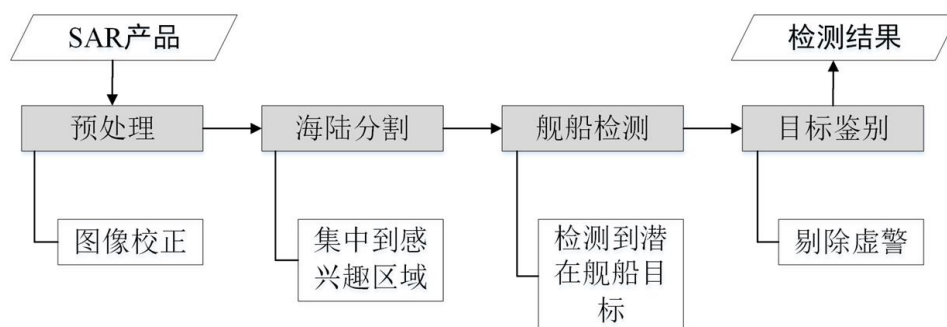


图 1-2 传统 SAR 图像舰船检测方法流程图

Figure 1-2 Workflow of traditional SAR ship detection methods

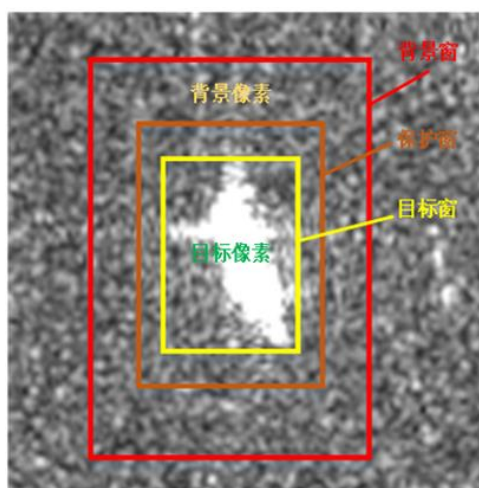


图 1-3 CFAR 方法的滑动窗口示意图

Figure 1-3 Sliding window in CFAR

随着硬件计算能力的提高和 SAR 图像数据量的不断增大，基于深度卷积神经网络的 SAR 舰船检测算法相继被提出。基于深度卷积神经网络的 SAR 舰船检测算法能够直接从原始影像数据中学习图像特征，算法在一个统一的端到端的框架下进行参数优化。因此，这类算法极大地提高了 SAR 图像中舰船检测算法的智能化程度，能够实现更加全面的图像信息表征，从而有效克服了传统算法所带来的缺陷。由于深度卷积神经网络的独特优势，近年来，在计算机视觉领域中发展出了多种基于深度卷积神经网络的目标检测算法。概括地讲，这些算法共包括两大类：一类是以 R-CNN（Region-based Convolutional Neural Network）系列为代表的基于候选框提取的双网检测方法；另一类是以 YOLO（You Only Look Once）系列为代表的基于边界框回归的单网检测方法。为了

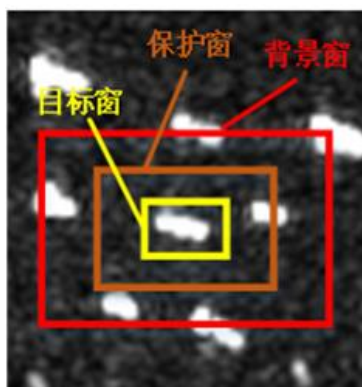


图 1-4 密集排布目标对 CFAR 窗口的影响

Figure 1-4 Effects of densely clustered targets for sliding windows in CFAR

提高 SAR 图像中舰船目标检测的性能, 学者们也基于深度神经网络方法, 相继研究了针对 SAR 图像舰船检测中特定问题的检测器。文献 [54] 提出了一种融合了最稳极限区域 (Maximum Stability Extremal Region, MSER) 方法的深度卷积网络结构用于高分辨率 SAR 图像中的舰船目标检测。国防科技大学计科锋教授课题组针对 SAR 图像中舰船目标的检测做了大量工作^[55-57]。其中, 文献 [56] 提出了一种改进的 Faster R-CNN 方法^[58], 融合了上下文与底层特征图用于 SAR 图像中的舰船目标检测。同年, 文献 [57] 在舰船检测方法中结合了传统的基于 CFAR 的方法和 Faster R-CNN 方法。Jiao 等在文献 [59] 中提出了一种沿用了 Faster R-CNN 框架的基于密集连接网络结构和多尺度金字塔结构的 SAR 图像舰船目标检测方法。Zhang 等提出了一种基于卷积神经网络的深度可分离快速 SAR 图像舰船目标检测方法^[60]。文献 [61-64] 在 SAR 图像舰船检测任务中, 融合了基于多尺度金字塔结构的卷积神经网络方法与视觉关注机制。中科院电子学研究所提出了一种基于海杂波分布和卷积神经网络的 SAR 图像舰船目标检测方法^[65], 该方法首先采用全卷积网络进行海陆分割, 然后通过分析高分三号图像中的海杂波分布情况, 从 K 分布、Gamma 分布和 Rayleigh 分布中选择基于 CFAR 方法的海杂波模型, 最后采用卷积神经网络方法对舰船目标进行鉴别。中科院遥感所王超研究员团队提出了一种结合 Single Shot MultiBox Detector (SSD)^[66] 和迁移学习的舰船检测方法^[67], 并且在我国高分三号数据上验证了该方法的有效性。西安电子科技大学杜兰教授团队在 SAR 图像数据扩充的基础上, 也提出了一种基于 SSD 网络结构的迁移学习方法用于 SAR 图像目标检测^[68], 杜兰等针对 SAR 图像中训练样本不足的情况, 利用已有的完备数据集来辅助检测场景复杂且训练样本不足的数据集, 并且在 MSTAR 和 MiniSAR 车辆数据集上验证了该方法的有效性^[69]。此外, 西安电子科技大学的焦李成教授团队综合考虑了视觉注

意模型与卷积神经网络方法,并采用多尺度金字塔模型提高对复杂背景以及多尺度舰船目标的检测能力^[70]。之后,该团队又在文献 [71] 中提出了一种基于自适应重校正机制的深度神经网络方法用于 SAR 图像多尺度与方向性舰船检测。该团队又于同年,在文献 [72] 中提出了一种称为 RADet 的方向性 SAR 图像舰船检测方法,该方法综合考虑了特征金字塔网络 FPN (Feature Pyramid Networks) 与多层视觉注意模型。复旦大学徐丰教授团队也在 SAR 图像舰船检测工作中做了大量研究,包括 [73-76]。针对极化 SAR 图像中的舰船目标检测问题,学者们也提出了各种基于深度卷积神经网络的方法,使得检测性能较传统方法得到了极大提升^[77, 78]。

1.2.4 极化 SAR 图像散射特性

极化 SAR 图像散射机制的研究主要在目标分解部分,其目的是将极化 SAR 图像中目标的散射过程分解为表示不同散射机理的若干项,每一项对应不同的物理意义。极化 SAR 散射机制具有明确的物理意义,且极化散射机制是图像分类与识别、目标检测等极化 SAR 应用的基本元素与重要基础。

极化 SAR 图像因其数据中蕴含着非常丰富的极化信息而得到各国研究者的广泛关注。这些极化信息对目标的物理特性、几何结构等特征具有较强的表征能力。然而,由于极化 SAR 图像独特的复数矩阵描述方式,导致其直观可视性非常差,大量蕴含在其中的信息不能被人类肉眼所直观观测到。为了解决这一问题,研究者们开展了大量的研究工作,并取得了一定的成果。其中,大多数为图像编码与基于目标极化分解的方法。

全极化 SAR 图像共包含 HH、HV、VH 和 VV 四个通道的矩阵存储复数据。其中,HH 极化表示雷达系统发射水平极化 (Horizontally Polarized) 电磁波,同时接收水平极化电磁波的过程;HV 极化表示雷达系统发射水平极化电磁波,同时接收垂直极化 (Vertically Polarized) 电磁波的过程;VH 极化和 VV 极化同理。目前有两类方法可实现 SAR 图像的直观可视化。第一类极化 SAR 可视化方法是基于图像编码的方法。这类方法是最简单也是最初被提出的图像伪彩色编码的方式。具体来讲,就是将 HH、HV 和 VV 三种极化方式图像的幅度值分别作为 RGB 颜色空间中的 R、G 和 B 三原色分量并进行归一化之后显示出来的伪彩色图。在这种方式可视化出来的伪彩色图像中,幅值差异大的目标能够很容易被区分开来,但是该方法不能对蕴含在图像中的丰富的极化信息进行有效显示与解读。第二类方法是基于极化目标分解理论的可视化方法。该类方法首先从原始 SAR 图像中提取具有物理意义的极化特征分量,然后采用 RGB 伪彩色图的方式将其可视化。在这类方法中,Pauli 分解是最初被提出并应用的经典且有效的模型之一^[79]。该方法基于 Pauli 基将散射过程分解为:各向同性的奇次散射,主要对平坦的地球表面、海上区域、球体、三角反射器等;各向同性的偶次散射,这种散射类型所对

应的典型散射体包括方位角是零度的二面角结构；与水平方向构成 45° 方向角的二面角结构。在满足互易定理的条件下，Pauli 基通常被简化成三种成分。这三种极化散射方式分别作为 RGB 颜色空间中的三个分量，以伪彩色图的形式合成并显示。Krogager 分解^[80] 通过一定的矩阵分解方法将对称散射矩阵分解成非相互正交的表示一定特殊目标特性的基，最终将极化散射矩阵分解为二面角散射、表面散射、螺旋体散射三个具有不同物理意义的分量。Krogager 分解方法充分利用了极化 SAR 图像中的相干特性，对高分辨率极化 SAR 图像较为适用。依据地球观测目标基不变的特性，Cameron 分解^[81] 首先将极化散射矩阵分解为互易分量 (reciprocal) 和非互易分量 (non-reciprocal)，之后互易分量再被细分为最大对称分量和最小对称分量。学者们通常通过最大对称分量变换对地物观测类型进行划分。这种极化目标分解方法着重强调在 Poincare 球上具有线性本征极化作用的对称目标。Cameron 分解方法是从散射矩阵着手分析，由于最初没有对目标进行相干性散射检验，因此并不适用于非相干散射区域。针对 Cameron 分解的这些缺陷，Touzi 在二面角和三面角正交基矢量空间中通过三个不同的参数以及二面角与三面角之间的相位差来表示 Cameron 分解中的最大对称分量^[82]。而且，该分解方法将最大对称分量细化为比 Cameron 分解更加完备的分量，使得该方法对目标中的对称散射取得较好的分割效果。在 Huynen 分解^[83] 中，研究者首先从平均 Mueller 矩阵中提取出一个平均单一目标 (Single Target) 和一个残留分量，然后采用一个 3×3 的矩阵描述平均单一目标。这个 3×3 像素大小的矩阵中，每个元素表示单一目标的不同物理特征，研究者们通常根据需求从中选取感兴趣的三个元素，从而合成可以直观显示的伪彩色编码图像。Huynen 首次提出将极化目标分解与散射机制相关联，从此在该方向上开展了大量科学研究。以上极化目标分解方法都是从极化散射矩阵进行分解，称为相干分解法。后来，学者们对二次相干矩阵和协方差矩阵进行分解，发展了一系列非相干极化目标分解方法。其中，Van Zyl 分解^[84] 也是一种基于 Mueller 矩阵的分解方法，作者将协方差矩阵分解为奇次散射、漫散射和偶次散射三个分量。除此之外，最著名的且应用最为广泛的三种极化目标分解方法为 Freeman-Durden 分解^[85]、Yamaguchi 分解^[86, 87] 和 Cloude-Pottier 的 $H/\alpha/A$ 分解^[88]。Freeman-Durden 分解构造了体散射、多次散射和 Bragg 散射，共三种典型的具有实际物理意义的目标散射模型。然后，以这三种散射模型作为矩阵分解的基，对协方差矩阵进行分解，最终得到能够有效描述雷达后向散射特性的三个分量。基于 Freeman-Durden 分解，Yamaguchi 考虑了在城镇等地物结构复杂区域的反射不对称现象（即 $S_{HH}S_{HV}^* \neq S_{VV}S_{HV}^* \neq 0$ ），提出了一个包含体散射、多次散射、表面散射和螺旋体散射的四分量散射模型。其中，螺旋体散射能够有效描述非均匀区域。Cloude-Pottier 的 $H/\alpha/A$ 分解基于相干矩阵的特征分解模型提取出极化熵、极化散射参数角、极化散射异向性值三个分量。在该模型中，极化熵用于描述雷达目标的极化程度，极化散射参数

角用于描述散射机理,极化异向性值用于判定区域散射类型。同样基于相干矩阵的特征分解模型, Holm 分解^[89] 将相干矩阵分解为纯目标、混合目标和非极化混合目标分量。但是在具有较大极化方位角的建筑区,其极化目标分解结果中存在散射类型与林区等植被覆盖区域有着严重的相似性。因此,针对极化目标分解问题,学者们也对城区复杂场景下的体散射过估计现象进行了广泛研究^[90-95]。一般来说,相干分解方法适用于地球观测中的纯目标,其回波信号具有相干性,可通过 Sinclair 散射矩阵描述;而非相干分解方法适用于处理分布式目标,这类目标可通过相干矩阵和协方差矩阵进行表征。

1.3 研究内容与组织结构

本课题重点围绕国家自然科学基金重点项目“高分辨率目标认知理论与关键技术”和十三五预研项目“基于脑机协同的**目标检测技术”等课题展开深入研究。近些年来,以卷积神经网络为代表的深度学习方法在多个 SAR 和极化 SAR 图像数据集上得到了有效验证,也在很多应用研究上取得了突破性进展。SAR 图像解译研究已经逐渐由传统的统计模型方法过渡到了与卷积神经网络等机器学习方法相结合的研究领域中。本课题以大数据背景下的 SAR 图像解译技术为契机,结合卷积神经网络与高斯过程等机器学习理论,研究 SAR 和极化 SAR 图像的目标/地物精细化解译方法,重点研究面向 SAR 图像舰船检测的目标精细化解译、多种极化方式下 SAR 图像的物理散射类型提取、多种探测波段下 SAR 图像的物理散射类型提取、SAR 图像精细化解译测试库构建与质量评估。论文主体部分组织结构和各部分之间的关联如 1-5 图所示。各章节的详细内容如下:

第二章是全文的基础知识部分,是全文的理论根基。详细介绍了关于 SAR 图像的四个方面内容。SAR 的成像基础介绍 SAR 图像的形成机理、分辨率特性和电磁波特性;接下来,从 SAR 图像的几何畸变和相干斑噪声两方面介绍了 SAR 图像的特点及其给 SAR 图像解译带来的挑战;在极化 SAR 理论方面,首先介绍了雷达极化的原理与极化 SAR 的某些应用点;另一方面,也详细介绍了极化 SAR 应用当中的非相干表征和相干表征方法,这是极化 SAR 图像解译的基础与出发点。

第三章重点研究基于卷积神经网络方法的 SAR 图像舰船目标检测,从而实现面向 SAR 图像舰船检测的目标精细化解译。本章在分析当前主流的基于卷积神经网络的 SAR 图像舰船检测方法的基础上,首先采用多尺度卷积神经网络方法同时实现了远海海域和近港/岸区域的舰船目标高精度检测;其次,针对检测框精确度不足、目标密集排布区域漏检严重、少量陆地虚警问题,采用级联卷积神经网络和基于脉冲余弦变换的检测方法联合实现 SAR 图像中舰船目标的精细化提取,并分析讨论了方位向模糊对舰船目标检测带来的影响。

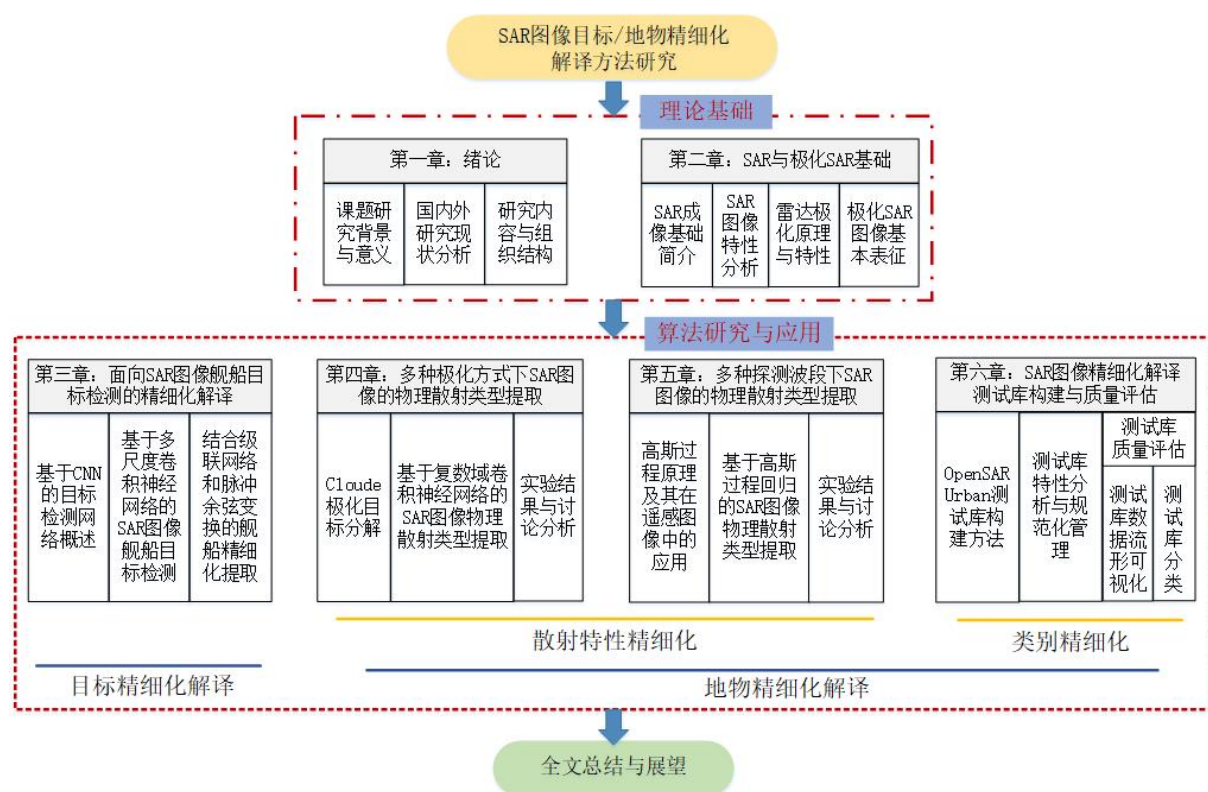


图 1-5 论文研究内容与组织架构

Figure 1-5 Contents of this thesis and their organizations

第四章重点从雷达极化的角度分析极化 SAR 图像中物理散射机制问题。全极化 SAR 图像对地物目标的物理散射特性具有较强的表征能力，而单极化和双极化 SAR 图像在该方面的表征能力非常有限。本章主要通过基于复数域卷积神经网络的方法实现从单极化和双极化 SAR 图像中提取具有物理意义的散射类型，并在此基础上探讨多种不同极化方式 SAR 图像的物理散射特性。基于复数域卷积神经网络的方法能够直接从原始复散射矩阵中学习散射参数，损失函数采用基础损失项结合对比约束项的方法实现。相关实验在德国宇航中心的多景 L 波段高分辨率 F-SAR 数据上得到了验证。

第五章实现面向多种探测波段 SAR 图像物理散射特性的精细化解译。本章的研究目标是从多波段 SAR 图像中提取具有物理意义的散射类型，通过基于核函数和平均期望的高斯过程回归方法实现跨雷达波段 SAR 图像中极化熵和极化散射角等参数预测，并在此基础上根据 Cloude 划分准则推导出图像中具有物理意义的散射类型（包括多次散射、体散射和表面散射），从而实现多种探测波段下 SAR 图像的物理散射类型提取。实验在 L 波段 F-SAR 数据、X 波段 TerraSAR 数据，以及 C 波段 Sentinel-1 数据上得到

了有效验证。

第六章贡献了一个基于中低分辨率 Sentinel-1 影像的大规模 SAR 图像地物分类测试库,实现了面向 SAR 图像地物精细化解译的测试库构建与质量评估。从遥感领域当中的测试库发展概况与目前的研究缺陷出发,在预定义的精细化城区两级分类体系基础上,重点介绍了 OpenSARUrban 测试库的构建方法和测试库的质量评估。在 OpenSARUrban 测试库构建方法中,详细介绍了原始 Sentinel-1 影像的收集和预处理方法、测试库类别体系的定义、测试库构建的关键步骤与方法、大规模测试库的组织形式、OpenSARUrban 测试库特性,以及测试库的规范化管理。在测试库质量评估方面,本文从基于快速压缩距离 (Fast Compression Distance, FCD) 和 t-SNE 的数据流形可视化以及基于典型特征提取和经典卷积神经网络方法的图像分类角度对测试库进行质量评估。

在全文总结部分,本文根据论文的研究内容进行总结归纳,给出本文的创新点。除此之外,本文也分析了当前 SAR 图像解译中的难点、分析论文研究的不足之处,并指出待解决的问题以及对未来研究的展望。

第二章 SAR 与极化 SAR 基础

2.1 引言

SAR 图像的主动式微波成像的原理导致其与光学图像有着本质的区别。SAR 的主动式发射与接收电磁波成像的特性导致所产生的图像不符合人类视觉系统的观测机制, 因此 SAR 图像的理解极具挑战性, 在图像上所呈现出的特点表现为乘性相干斑噪声和几何畸变等现象。本章介绍 SAR 与极化 SAR 的基础知识, 具体的内容包括 SAR 成像基础、由于主动式微波成像导致的 SAR 图像独有的特性; 此外, 本章针对极化 SAR 图像的解读, 重点介绍了极化 SAR 的原理与特性、极化 SAR 图像特性和基本表征的相关内容。

2.2 SAR 成像基础简介

雷达, 源于英文单词 RADAR 的音译, 是 RAdio Detection And Ranging 的缩写, 意思是无线电探测与测距¹。从英文释义中即可理解, 雷达是一种主动式采用微波成像的设备, 通过无线电定位的方法发现目标并测定其位置。雷达图像依靠加载在卫星 (星载) 或飞行器 (机载) 上的天线发射和接收无线电电磁波信号, 并通过一定的成像技术获取。当雷达设备发射电磁波时, 电磁波与地球表面接触后会朝着不同的方向散射回波, 部分回波信号被雷达天线接收到, 继而通过雷达成像系统, 根据回波信号的强度和方向记录地球表面物体的散射特性。图 2-1 采用一种易于理解的方式展示了 SAR 系统的成像几何关系。图中, P_1 & P_2 是搭载雷达的卫星和飞行器在近似匀速运动过程中的两个位置点, R 是飞行过程中雷达天线与地球表面物体之间的距离, 在侧视成像关系中称作斜距, R_1 & R_2 是零多普勒频率校准后的斜距。

早期的雷达系统主要用来解决有无目标的问题, 返回的信息并不是图像, 而是回波脉冲或回波的强弱。进入到 20 世纪 50 年代, 研究者们采用多普勒锐化技术^[96-98] 实现了雷达回波方位向的高分辨率成像, 后来才逐渐发展成了 SAR。突破了原始的雷达概念, SAR 的研制成功使得雷达具备了对目标和区域进行成像并区别的能力, 这些进展能够为本文提供更多的地物目标细节信息。为了获取对地观测的二维图像, SAR 成像技术在距离向上采用线性调频技术^[99-101] 将相邻或相近的物体从雷达回波中分离开来; 在方位向上, SAR 通过平台运动构造了一个虚拟的大孔径天线, 从而实现了方位向上的高分辨

¹<https://en.wikipedia.org/wiki/Radar>

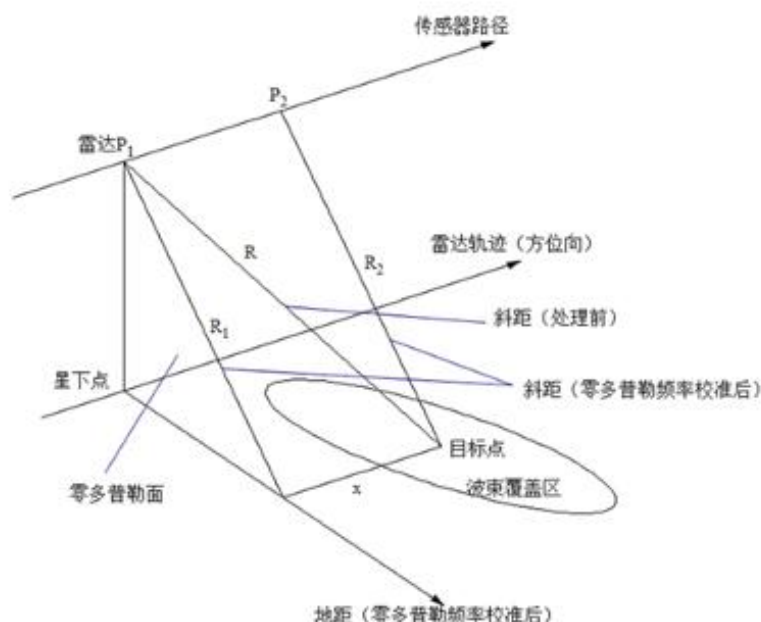


图 2-1 SAR 成像几何关系

Figure 2-1 Geometric relationships of SAR image acquisition

率成像。

在雷达成像的过程中，当雷达天线固定不动时，接收天线往往只能接收到一小部分后向散射信号；而当雷达天线快速前进时，接收天线能够收集到从地表后向散射到各个方向的电磁波信号。雷达合成孔径的原理就是利用天线的运动，将小孔径的雷达天线虚拟成一个较大孔径的雷达天线，从而使 SAR 成像达到类似大孔径天线的探测效果。图 2-2 展示了合成孔径的原理。图中虚线粗略表示雷达天线运动的运动轨迹， β 为合成孔径夹角， L_s 表示方位向覆盖范围。

2.3 SAR 图像特性

雷达散射回波信号通过一定的成像处理方式将其转换成地面二维散射信息，即人类视觉可见的图像。通常来讲，SAR 成像处理后得到的数据可以被认为是地面后向散射的二维分布，这些数据是按照复数据的形式保存的，也就是每个地面分辨率单元的散射回波信号都是由实部和虚部共同构成的。事实上，每个雷达图像都可以看作是存储地面后向散射的二维矩阵。与光学图像相比，SAR 图像具有以下特性：

- 由于云雨雾等天气因素对雷达成像的影响较弱，SAR 图像能够实现全天候作业；

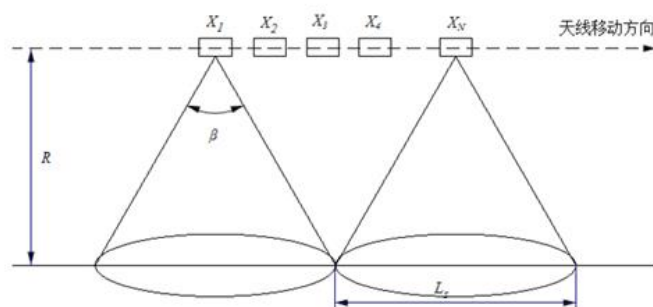


图 2-2 合成孔径原理

Figure 2-2 Synthetic aperture theory

- 由于 SAR 系统主动发射与接收电磁波的性能与白昼无关，SAR 图像能够实现全天时作业；
- 雷达电磁波在传播过程中，几乎不受大气成分影响，因此 SAR 图像适合多时相分析，如地物覆盖、海上区域等的动态监视；
- SAR 图像对地表粗糙度、材质的介电常数等因素较敏感，因此 SAR 图像中不同的地物目标具备不同的散射强度与相位；
- 由于电磁波的表面穿透特性受电磁波波长的影响较大，因此不同工作波段的 SAR 系统所采集的数据信息不同；如波长较长的电磁波能够穿透植被和树叶，因而图像中所反映出的是树干或者地球表面的散射回波，而短波长的电磁波由于穿透性较弱，其采集的图像反映的是叶簇的散射回波；
- 雷达图像的分辨率不受雷达侧视角、雷达系统工作波长，以及搭载传感器的平台飞行高度的影响。

然而，由于 SAR 图像的主动式成像和相干成像的原理，SAR 图像并不符合人眼视觉系统的观测机制，SAR 图像的理解也变得异常困难。

2.3.1 几何畸变

与光学图像相比，由于 SAR 成像过程中采用斜侧视成像的方式，加上成像过程受到地球表面起伏的影响较为严重，因此，SAR 图像具有严重的几何畸变特征。这种现象在高分辨率图像中的城市等人造结构密集的地区尤其明显。SAR 图像的几何畸变现象^[102, 103] 具体来说主要表现在叠掩 (layover)、二次散射 (double bounce)、屋顶散射 (roof)、阴影 (shadow) 等 (如图 2-3 所示)。叠掩是指在同一时刻雷达接收到的回波中同时混有屋顶散射回波、立面墙体散射回波和地表散射回波中的两种或两种以上散射的现象。二

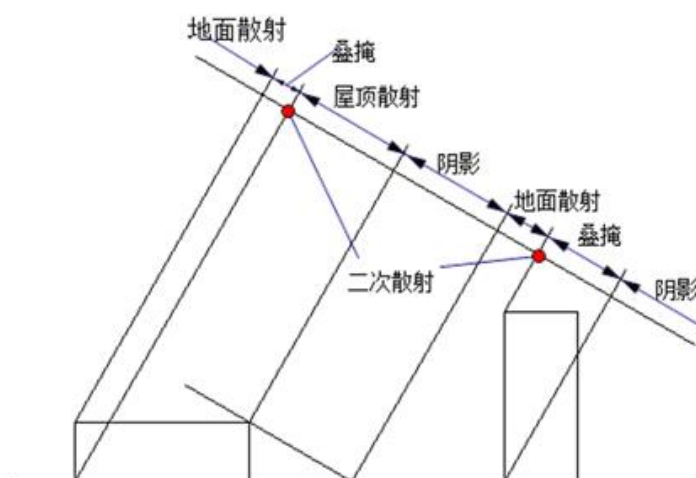


图 2-3 SAR 图像几何畸变现象

Figure 2-3 Geometric distortions in SAR images

次散射是指在二面体或三面体位置处，雷达回波信号强度远远大于其他位置处，在 SAR 图像上表现为一条非常明显的高亮线型或 L 型结构。屋顶散射是指雷达天线只接收到来自屋顶的散射回波的现象。阴影是指由于其他地面物体的遮挡，雷达发射的电磁波所照射不到的区域在图像上的表现。地表散射是指雷达天线只接收到了来自地面的散射回波，而无其它成分的现象。这些由于侧视成像而带来的几何畸变的现象也在很大程度上增加了图像理解的难度。但是，合理利用这些几何特性也能够为 SAR 图像解译提供重要信息。

2.3.2 相干斑噪声

SAR 图像中的乘性相干斑噪声是由 SAR 成像过程中的相干成像原理造成的^[104, 105]。雷达系统在通过电磁波照射地球表面的过程中，其接收到的合成散射回波并不完全由散射回波系数决定。由于各个散射点的回波是相干的，这些回波相互干涉，导致合成的电磁波矢量的振幅和相位在原有的基础上存在一定的起伏，这就是相干斑噪声的形成原理。在实际 SAR 图像中，相干斑噪声是以一种乘性噪声的形式存在的，表现在图像上为图像灰度的剧烈变化。正是受到这种乘性相干斑噪声的影响，SAR 图像无法精确反映地物目标的散射特性，使得图像的精细结构隐藏在所获得的 SAR 图像中，从而增加了研究人员从 SAR 图像中提取目标信息的难度。

光学与 SAR 图像的对比示例如图 2-4 所示。从这两幅图中可以看到，与光学图像相比，SAR 图像受到几何畸变和相干斑噪声的影响非常严重，这正是由 SAR 的斜侧视成

像与相干成像原理导致的。

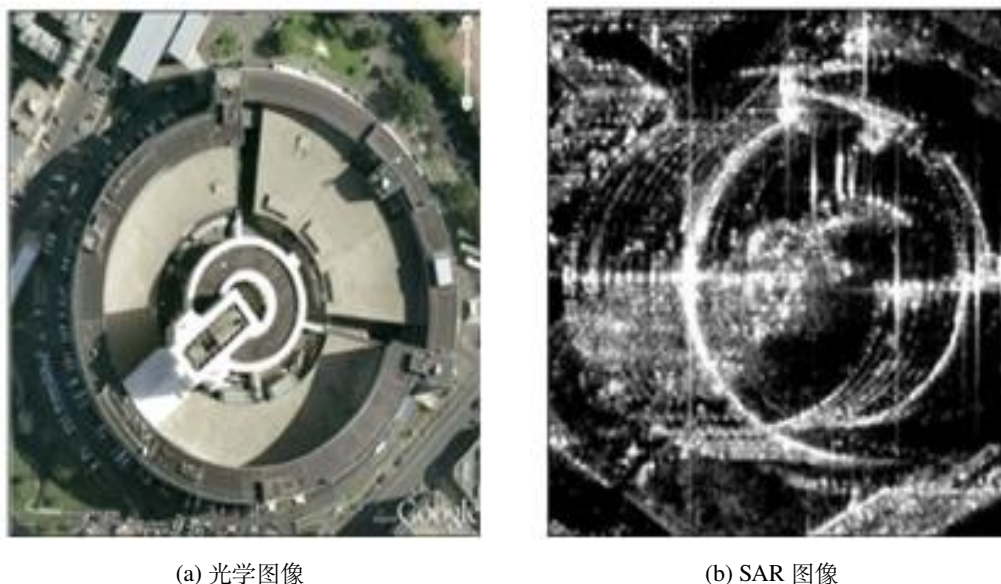


图 2-4 光学图像与 SAR 图像对比

Figure 2-4 Comparison of an optical and SAR image example

2.4 雷达极化原理与特性

电磁波的极化方向指电场矢量的端点随时间变化而形成的轨迹，也就是电磁波传播过程中电场矢量的瞬时方向。在电磁波传播过程中，当电场矢量方向在垂直于电磁波传播方向平面内的投影是一条线段时，这种类型的电磁波被称为线极化波。一般来说，电磁波的极化方式分为水平极化和垂直极化两种方式。当 SAR 系统发射和接收的电磁波都是水平极化波，则获得的影像为水平极化影像 (HH)；当 SAR 系统发射与接收的电磁波分别为水平极化波和垂直极化波时，相应的 SAR 图像为 HV 极化方式图像。同样的道理，研究人员根据雷达系统发射与接收电磁波的极化方式不同，有 VH 极化和 VV 极化的图像。HH 和 VV 极化方式的 SAR 图像称为同极化 (Co-Polarized) SAR 图像；HV 和 VH 极化方式的 SAR 图像称作交叉极化 (Cross-Polarized) SAR 图像。具有双极化和全极化成像能力的 SAR 系统一般能够同时获取同一成像区域下两种或者四种不同极化方式的 SAR 图像，且图像均以复数据的形式保存。

通常情况下，与同极化 SAR 图像相比，交叉极化方式 SAR 图像的散射回波信号较弱。鉴于此，星载成像 SAR 系统通常采用像 HH 极化和 VV 极化这样的同极化方式。对大多数地物目标而言，HH 极化方式 SAR 图像能够获得的回波信号最强，地物目标后向

散射之间具有较大差异或对比度。因此，在实际情况中，大多数单极化 SAR 系统都采用 HH 极化方式来获取对地观测图像。

雷达发射与接收电磁波波长在雷达系统中具有极其重要的关键作用。根据雷达发射与接收电磁波波长的差异，SAR 和极化 SAR 图像可分为不同的波段，如 L 波段、C 波段、X 波段、P 波段等。长波波长的 SAR 图像具有分辨率较低和穿透能力更强的特点，相反地，短波波长的 SAR 图像则分辨率更高、穿透能力较弱、并且更容易被吸收。对于森林地区来说，该区域在长波段 SAR 图像上体现的大多是地面目标或者树干等的后向散射，而在短波段 SAR 图像上体现的大多是叶簇或者树枝等结构的后向散射。图 2-5 综合反映了不同波段的雷达电磁波的后向散射机理（图中 5-8 示例）以及各种散射类型，如二次散射、体散射、表面散射（图中 1-4 散射示例）。

表 2-1 雷达电磁波波段比较

Table 2-1 Radar electro-magnetic bands comparison

波段代号	标称波长/cm	频率范围/GHz	波长范围/cm
P	/	0.23-1	130-30
L	22	1-2	30-15
S	10	2-4	15-7.5
C	5	4-8	7.5-3.75
X	3	8-12	3.75-2.5
Ku	2	12-18	2.5-1.67
K	1.25	18-27	1.67-1.11
Ka	0.8	27-40	1.11-0.75
U	0.6	40-60	0.75-0.5
V	0.4	60-80	0.5-0.375
W	0.3	80-100	0.375-0.3-30

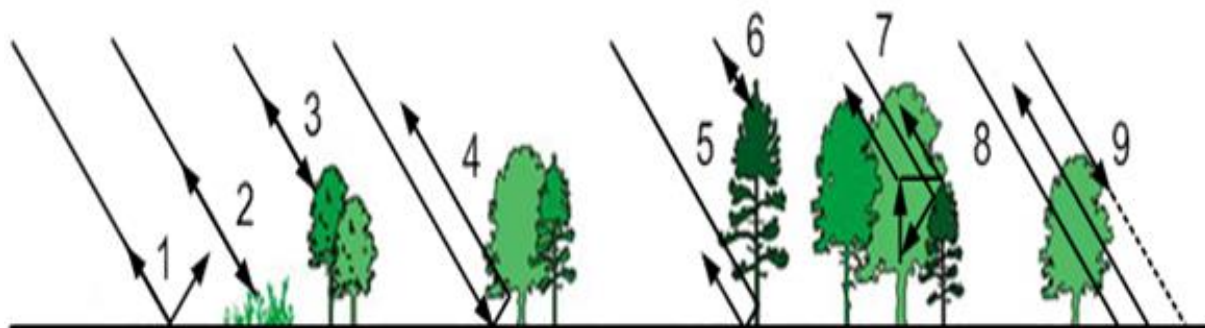


图 2-5 极化 SAR 图像后向散射特性

Figure 2-5 Backscattering characteristics in PolSAR images

2.5 极化 SAR 图像基本表征

2.5.1 极化散射矩阵及其向量化表征

一般地，雷达发射与接收电场矢量的关系在数学上可以表示为

$$\begin{aligned}
 E^{\text{bk}} &= \begin{bmatrix} E_{\text{H}}^{\text{bk}} \\ E_{\text{V}}^{\text{bk}} \end{bmatrix} = \frac{\exp(jk_0 R)}{R} \cdot S \cdot E^{\text{in}} \\
 &= \frac{\exp(jk_0 R)}{R} \cdot \begin{bmatrix} S_{\text{HH}} & S_{\text{HV}} \\ S_{\text{VH}} & S_{\text{VV}} \end{bmatrix} \cdot \begin{bmatrix} E_{\text{H}}^{\text{in}} \\ E_{\text{V}}^{\text{in}} \end{bmatrix}
 \end{aligned} \quad (2-1)$$

其中， $S = \begin{bmatrix} S_{\text{HH}} & S_{\text{HV}} \\ S_{\text{VH}} & S_{\text{VV}} \end{bmatrix}$ 是雷达散射矩阵， $(\cdot)_{\text{H}}$ 和 $(\cdot)_{\text{V}}$ 分别表示水平和垂直极化分量， E^{bk} 和 E^{in} 分别表示散射回波与发射电磁波中的电场矢量， R 表示雷达天线与地物目标之间的距离， k_0 为雷达波数。雷达极化散射矩阵 S 是用于表示一个雷达目标或者一个雷达目标群多种极化响应的一种基本表达方式。极化散射矩阵通过多维复数据的表达形式完整描述了地物目标的散射特性。

极化 SAR 的四个极化通道数据经过定标处理和数据压缩后生成单视复数 (Single-Look Complex, SLC) SAR 图像。在满足后向散射对称的互易定理条件下，两种交叉极化的回波图像之间可以有 $S_{\text{HV}} = S_{\text{VH}}$ 的关系。雷达的散射矩阵将雷达后向散射的强度、相位和极化信息统一表示与记录，能够完整描述地物目标的电磁散射特性。以单视复数据形式记录的散射矩阵不仅与地面目标本身的形状、几何结构、材料属性、介电性等因素有关，而且与雷达与地面目标之间的相对位置关系、姿态、雷达的工作频段等因素有关。

在满足互易定理的条件下，基于 Borgeaud 基的矢量化散射矩阵表示为

$$\vec{k}_B = \begin{bmatrix} S_{HH} \\ \sqrt{2}S_{HV} \\ S_{VV} \end{bmatrix} \quad (2-2)$$

其中， \vec{k}_B 称为常规散射矢量，该矢量以一种非常简单直观的方式涵盖了极化散射矩阵中的各个元素。

此外，还存在另外一种更为实用的完全正交矩阵基为 Pauli 基的向量表达方式。这种表达方式的数学定义为

$$\vec{k}_P = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{HH} + S_{VV} \\ S_{HH} - S_{VV} \\ 2S_{HV} \end{bmatrix} \quad (2-3)$$

其中， \vec{k}_P 表示 Pauli 基向量。

2.5.2 相干矩阵与协方差矩阵

通常情况下，极化散射矩阵用于地物目标是确定的点目标情况下。然而，在现实情况中，大多数目标都是分布式目标，SAR 图像需要滤波处理以降低相干斑噪声对图像的影响。这时候，就需要用到图像的二阶统计特性。常用的二阶统计表征方式包括相干矩阵和协方差矩阵，这两种表达方式都包含了雷达对地观测所得到的全部极化信息。

基于 Borgeaud 基的极化协方差矩阵定义为

$$C = E\{\vec{k}_B \cdot \vec{k}_B^\dagger\} = \begin{bmatrix} E\{|S_{HH}|^2\} & \sqrt{2}E\{S_{HH}S_{HV}^*\} & E\{S_{HH}S_{VV}^*\} \\ \sqrt{2}E\{S_{HV}S_{HH}^*\} & 2E\{|S_{HV}|^2\} & \sqrt{2}E\{S_{HV}S_{VV}^*\} \\ E\{S_{VV}S_{HH}^*\} & \sqrt{2}E\{S_{VV}S_{HV}^*\} & E\{|S_{VV}|^2\} \end{bmatrix} \quad (2-4)$$

其中， $E\{\cdot\}$ 表示平均期望运算， $|\cdot|$ 表示求散射回波的幅度运算， $(\cdot)^*$ 表示复共轭运算， $(\cdot)^\dagger$ 表示向量或矩阵的共轭转置。

同理，基于 Pauli 基的极化相干矩阵可表示为

$$T = E\{\vec{k}_P \cdot \vec{k}_P^\dagger\} = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{12}^* & T_{22} & T_{23} \\ T_{13}^* & T_{23}^* & T_{33} \end{bmatrix} \quad (2-5)$$

其中，

$$T_{11} = \frac{1}{2}E\{(S_{HH} + S_{VV})(S_{HH} + S_{VV})^*\} \quad (2-6)$$

$$T_{12} = \frac{1}{2}E\{(S_{HH} + S_{VV})(S_{HH} - S_{VV})^*\} \quad (2-7)$$

$$T_{13} = E\{(S_{HH} + S_{VV})(S_{HV})^*\} \quad (2-8)$$

$$T_{22} = \frac{1}{2}E\{(S_{HH} - S_{VV})(S_{HH} - S_{VV})^*\} \quad (2-9)$$

$$T_{23} = E\{(S_{HH} - S_{VV})(S_{HV})^*\} \quad (2-10)$$

$$T_{33} = 2E\{(S_{HV})(S_{HV})^*\} \quad (2-11)$$

极化协方差矩阵与极化相干矩阵均为半正定 Hermitian 矩阵，二者具有相同的特征值，区别在于每个特征值所对应的特征向量不同。极化协方差矩阵与极化相干矩阵之间可以相互转换，在满足互易定理条件下，二者之间的转换关系为

$$T = A \cdot C \cdot A^{-1} \quad (2-12)$$

其中，

$$A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & \sqrt{2} & 0 \end{bmatrix} \quad (2-13)$$

2.6 本章小结

本章作为全文的理论基础，分别概述了 SAR 图像的成像原理、SAR 图像区别于一般光学图像的特性、雷达极化原理、以及极化 SAR 图像的基本表征方式。SAR 成像原理方面主要介绍了 SAR 的侧视成像几何关系和合成孔径原理。在 SAR 图像特性方面，主要从地物目标的几何畸变和相干斑噪声两方面展开。此外，基于雷达极化原理，详细介绍了相干和非相干的极化 SAR 图像基本表征方式。

第三章 面向 SAR 图像舰船目标检测的精细化解译

3.1 引言

SAR 图像舰船目标检测在军事和民用领域中都具有广泛的应用价值。在军事领域中,考虑到海军实力在国家综合军事力量中的重要作用,舰船目标的快速高精度检测,尤其是海上舰船等目标的实时动态监测,能够在军事作战中提供重要情报。在民用领域中,各国的海上船只活动日益频繁,海上货船、渔船等舰船目标的实时高精度检测能够在很大程度上预防海上交通事故,保障海上活动安全。

舰船目标检测的早期研究主要采用基于 CFAR 的方法,这种检测方法通过多阶段的算法过程实现。研究者们首先对图像进行海陆分割,剔除陆地等非感兴趣区域;然后通过统计建模方法从图像中提取可能存在舰船目标的区域,即检测候选区;接下来,通过对候选区域提取图像特征和分类器分类得到最终的检测结果,也就是目标鉴别的过程。这种多阶段的舰船目标检测方法虽然算法实现较为容易,但是仍然存在一些固有的缺陷:1) 多阶段算法实现过程繁杂且智能化程度不高;2) 人工设计的特征在面对高海况等一些复杂背景的情况下效果不够理想;3) 检测结果严重依赖背景建模与参数估计的准确性。

近年来,基于卷积神经网络等深度学习的目标检测方法层出不穷,其不仅能够克服传统方法的这些缺陷,也能够各种复杂背景和高海况条件下取得比较理想的检测效果。这类舰船目标检测方法通过一个端到端的网络结构与优化算法从原始图像中自动学习图像特征。由于整个过程是在一个一体化的检测框架中自动计算,在很大程度上缓解了传统 SAR 图像舰船检测方法的固有缺陷,实现了较为智能化的目标检测。

本章主要实现面向 SAR 图像舰船目标检测的精细化解译。首先,本章通过一个改进的多尺度卷积神经网络同时实现了远海海域和靠岸/港等舰船目标密集排布区域的高精度检测。其次,为了使得检测框更加紧凑精确,并减少密集排布目标漏检和陆地虚警,本章提出了一种结合级联卷积神经网络和频域脉冲余弦变换的检测方法,并借助航天飞机雷达地形测绘任务 (Shuttle Radar Topography Mission, SRTM) 数字高程模型 (Digital Elevation Model, DEM) 数据¹剔除大部分陆地虚警,最终实现 SAR 图像中舰船目标的精细化提取。

¹<http://srtm.csi.cgiar.org/>

3.2 基于卷积神经网络的目标检测方法概述

卷积神经网络方法不仅在自然图像目标检测中取得了令人满意的成果，在多种遥感图像目标检测的应用问题上也成绩斐然。基于卷积神经网络方法的 SAR 图像目标检测面临着新技术带来的机遇，同时也存在着诸多挑战。相比传统的机器学习方法，卷积神经网络的主要优势在于：1) 具有极强的特征学习能力，强调通过构建多层网络结构模型逐层学习线性与非线性特征变换，通过优化目标函数与反向传播机制更新网络参数，最终能够从海量的训练数据中学习具有更强区分能力和信息量更加丰富的特征表示，网络学习到的高层语义特征具有表征数据本质的内在含义的能力；2) 具有端到端学习的天然优势，模型直接作用于原始数据，整个模型训练过程直接优化预先设计的目标函数，而传统方法通过分阶段多步骤的检测方式，将整个目标检测过程分为几个独立的处理步骤，并不能在一个统一的框架下优化模型参数。正是由于这些优良的性质，卷积神经网络方法在短短几年内兴起，并能够引领人工智能研究的一大浪潮，在近景图像目标的分类识别、目标跟踪、自然语言处理等领域取得了突破性的研究进展。到目前为止，基于深度卷积神经网络的目标检测算法主要分为两大类，它们分别是基于区域候选框预测的双网检测算法和基于检测框回归的单网检测算法。一般来说，双网检测能够取得较好的检测精度，但是网络的检测效率略低。与之相反，单网检测方法在检测效率上具有较为明显优势，但是会损失检测精度。下面分别以 R-CNN 系列和 YOLO 系列为代表从双网检测算法和单网检测算法两方面展开介绍基于卷积神经网络的目标检测算法。

3.2.1 基于区域候选框预测的双网检测算法

区域候选框预测是预先找出图像中可能出现目标的位置，然后再通过特征提取与分类（目标区域与非目标区域）的方法确定最终的检测结果，后者也被称为目标鉴别的过程。但是，由于区域候选框预测过程中利用了图像的纹理、边缘等信息，可以保证在选取较少窗口的情况下保持较高的召回率。这大大降低了后续操作的时间复杂度，并且获取的候选窗口要比滑动窗口的质量更高（滑动窗口固定长宽比）。比较常用的候选框提取算法有选择性搜索方法（Selective Search）^[106] 和边缘检测框算法（Edge Boxes）^[107]。

作为深度学习领域目标检测任务中的开山之作，R-CNN 网络的检测过程可以概括为以下几个关键点：1) 输入测试图像，利用选择性搜索方法从图像中提取目标候选区域；2) 对每个候选区域采用图像缩放的方式变换到固定的尺寸；3) 将缩放后具有统一固定尺寸的图像输入到卷积神经网络中，并从最后一个全连接层提取其输出特征作为候选框的特征；4) 基于所提取到的网络深度特征，结合 SVM 分类器对候选框区域进行分类；5) 边框回归，即对候选区域进行线性回归校正，使得最终的检测框逼近真值窗口。R-CNN 网络充分利用了卷积神经网络和候选框提取的优势，但是该网络框架也存在诸

多缺陷：1) 网络训练采用微调网络、训练 SVM 分类器和边框回归器等多个阶段，步骤繁琐复杂；2) 网络训练耗时，计算机磁盘空间利用率低；3) 候选框缩放导致目标大小、尺寸比例等发生形变，从而影响检测结果。

针对 R-CNN 网络中需要对每个候选框图片提取深度特征从而影响目标检测效率的问题，空间金字塔池化网络 (Spatial Pyramid Pooling Network, SPP-NET)^[108] 首先对整张图片一次性提取深度特征，再将候选框区域映射到网络的最后一个卷积层。由于候选框的尺寸不同，映射到最后一个卷积层后的特征图尺寸不尽相同。针对这个问题，SPP-NET 通过空间金字塔池化的方式提取到相同维度的特征，然后再进行检测框的分类和回归。

Fast R-CNN^[109] 结合了 R-CNN 和 SPP-NET 的综合优势，并且在网络优化过程中引入多任务损失函数，使整个网络的训练和测试变得更加方便。在候选框提取上，仍然采用选择性搜索的方法，提取指定数量的候选框区域。与 SPP-NET 相同，Fast R-CNN 首先对整张图片一次性提取深度特征，然后将目标候选区域映射到最后一个卷积层，并采用空间金字塔池化操作得到维度一致的特征描述，网络的分类和边界框回归通过多任务损失函数优化的方式完成。

作为 Fast R-CNN 网络的改进，Faster R-CNN 网络的作者提出了一个 RPN 网络结构作为候选区域提取网络^[58]。这个子网络能够代替选择性搜索方法，在提取高质量候选框的同时也极大地减少了网络候选框提取的运算耗时，后续的目标鉴别流程沿用了 Fast R-CNN 中的目标鉴别方法。此外，Faster R-CNN 中 RPN 子网络与目标鉴别网络共享卷积特征提取层的机制大大降低了网络参数量。

3.2.2 基于检测框回归的单网检测算法

以 Faster R-CNN 系列为代表的双网目标检测方法虽然在检测精度上具有较大的优势，但是，其检测效率仍然不如单网检测。以 YOLO 系列为代表的单网目标检测方法直接从原始图像中回归出目标所在的位置及其属性。YOLO v1^[110] 检测器的主干网络采用了类似 GoogleNet 的结构。在检测过程中首先将原始图像划分为一个一个的小尺寸单元格，如果目标真值位置的中心点落在某一个单元格中，则该单元格将负责检测这个目标。对于每个候选框，网络同时预测了一个位置信息和一个置信度。其中，置信度是通过所属单元格是否包含目标和预测的 IoU 值计算得到。YOLO v1 网络检测目标的过程大致可以归纳为以下几点：1) 将给定的图像划分为 $S \times S$ 的网格；2) 在每个划分网格上预测两个候选检测框，因此会出现 $S \times S \times 2$ 个目标候选框；3) 基于预测的目标候选框，根据指定的阈值去除可能性较低的窗口，最后采用非极大值抑制 (Non-maximum Suppression, NMS) 操作去除冗余窗口。YOLO v1 的检测结果中容易出现定位错误，但

是该网络在区分前景与背景方面能力较强。另外, YOLO v1 网络对小尺寸目标和密集排布目标的检测效果并不理想。与 YOLO v1 网络相比, YOLO v2^[111] 网络主要在以下几个方面予以改进: 1) 通过在网络卷积层后采用批量归一化 (Batch Normalization, BN) 操作提高网络模型的鲁棒性, 缓解了数据偏差问题, 这在一定程度上起到了正则化的作用; 2) 该网络借鉴了 Faster R-CNN 中 RPN 网络的锚框思想, 通过增加候选框的个数提高检测结果的查全率; 3) 通过 k -means 聚类方法学习出锚点检测框, 并通过定义新的误差度量方式来解决目标尺寸大小对锚框的敏感性; 4) 通过跳跃层 (Passthrough Layer) 将来自底层的高分辨率特征和来自高层的低分辨率特征拼接起来, 从而提高对小尺寸目标和密集区域目标的检测效果; 5) 通过采用 DarkNet-19¹ 作为主干网络结构降低模型参数量, 从而提高模型训练效率。随后, YOLO v3 网络的作者在其所提出的网络 [112] 中将 DarkNet-53 作为主干网络结构。DarkNet-53 网络结构一方面采用全卷积网络结构减少了网络参数数量, 另一方面通过引入残差结构, 从而能够取得与 151 层残差网相当的效果, 并且大大减少了网络计算量。

3.3 基于多尺度卷积神经网络的 SAR 图像舰船目标检测

3.3.1 研究动机

Faster R-CNN 基于卷积神经网络方法通过 RPN 网络完成对目标候选框区域的提取, 并通过 Fast R-CNN 网络实现目标鉴别功能, 且 RPN 网络与 Fast R-CNN 网络共享深度网络特征提取层, 提升了目标检测的性能和速度。Faster R-CNN 的一个重要特点是在通过 RPN 网络提取候选框的过程中采用了锚框设计机制, 利用 CNN 的降采样特性, 在最后一层卷积特征映射上采用滑动窗的方式在每个位置上对参考窗进行分类和边界框回归, 取代了之前的在原始图像上采用滑动窗搜索的方法进行候选区域提取的策略。

基于 Faster R-CNN 的优良特性, 研究者们将该方法应用到 SAR 图像的舰船目标检测中, 并且取得了比传统的基于 CFAR 的检测方法更好的性能^[55, 56, 113]。但是, 在直接将 Faster R-CNN 方法应用到 SAR 图像舰船目标检测的过程中, 本文发现会存在一些不可避免的难点与挑战: 1) SAR 图像舰船目标的背景更加复杂, 高海况和复杂洋流等低信噪比区域的检测困难; 2) 在近岸/港等舰船目标密集排布的区域中, 难以实现准确有效的检测效果; 3) 由 SAR 图像的低分辨率特性和宽幅特性所带来的目标尺寸普遍较小的现象, 也给 SAR 图像中舰船目标检测增加了难度。已有的研究表明, 针对复杂背景问题, 采用 Faster R-CNN 这种端到端的深度卷积神经网络方法能够取得比传统多阶段方法更好的效果。但是, Faster R-CNN 网络难以在近岸/港等目标密集排布区域中取得

¹<https://pjreddie.com/darknet/>

较好的检测效果。本文通过采用多尺度卷积神经网络同时在 SAR 图像中远海海域和近岸/港区域实现较高精度的舰船目标检测效果。

3.3.2 方法详介

为了能够同时在远海海域和近岸/港区域实现高精度的舰船目标检测，本节采用一种改进的多尺度卷积神经网络方法。由于多尺度卷积神经网络方法^[114]沿用了 Faster R-CNN 的双网检测框架，仍然采用区域候选框提取网络与目标鉴别网络相结合的思路，保证了目标检测的高精度特性。针对 SAR 图像的较低分辨率特性和宽幅特性所造成的目标尺寸小以及目标密集排布检测困难的问题，本章采用了一种改进的多尺度卷积神经网络方法。该方法由被称作全面候选框提取网络（Exhausted Ship Proposal Network, ESPN）的子网络提取 SAR 图像中的疑似舰船目标区域，然后最终的检测结果通过精确目标鉴别网络（Accurate Ship Discrimination Network, ASDN）实现，并且二者通过主干网络框架耦合的方式实现特征共享。图 3-1 所示为改进的多尺度卷积神经网络的网络结构图，其中主干网络采用 VGG19 网络结构^[115]。

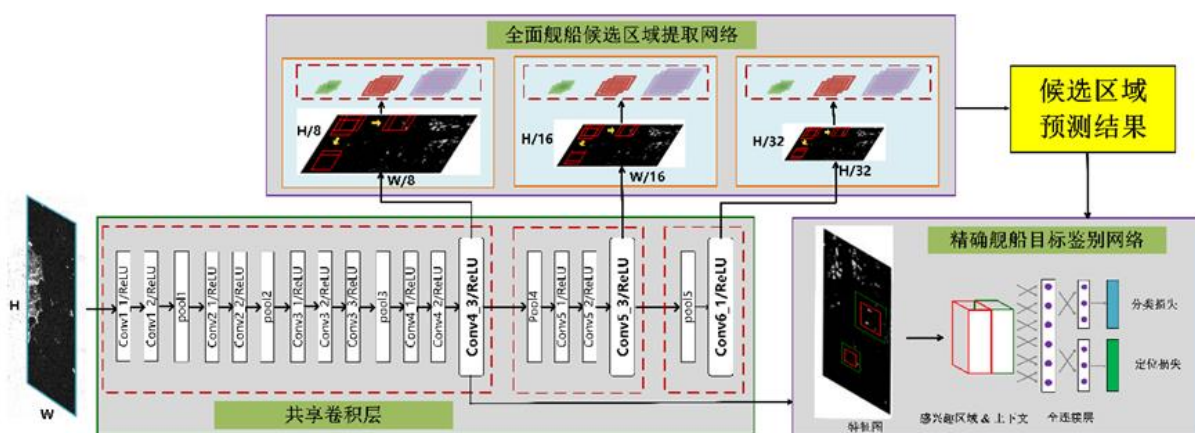


图 3-1 改进的多尺度卷积神经网络的网络结构图

Figure 3-1 Overview of the revised multiscale-CNN

为了同时实现在远海海域和近港/岸等舰船目标密集排布区域中的检测效果，在网络结构设计过程中，本文所采用的改进多尺度卷积神经网络主要遵循两个设计原则：

- 在区域候选框提取阶段本着减少漏检目标的原则，使得候选区域尽可能全面地覆盖舰船目标真值区域；
- 在舰船目标鉴别网络中本着减少虚警的原则，尽可能精确地实现 SAR 图像中的舰船目标和非舰船目标区域的鉴别，从而实现更加精确的目标检测。

由于 Faster R-CNN 中最后一个卷积层所输出的特征图尺寸小导致 RPN 网络对小尺寸目标的特征表达缺陷甚至出现信息缺失的现象。为了解决该问题,本文在 ESPN 网络中,借鉴多尺度卷积神经网络中多个典型网络层特征图相融合的策略,并设置合适大小和比例的锚框尺寸。此外,在每个典型特征层分支上,拓展出一系列不同大小的滑动窗滤波器组,通过这一系列小尺寸的滤波器组与所学习到的各个典型特征层输出的特征图进行卷积运算,从而能够使网络在候选区域预测位置处预测到比较完备的疑似舰船目标区域,最终通过候选框综合的方式提高对 SAR 图像中小尺寸且密集排布的舰船目标的区域预测能力。总的来说,ESPN 网络通过将每个拓展的典型网络子分支作为候选框提取通道,同时将每个通道提取的结果之并集作为最终的候选框区域,从而达到减少漏检的目的。表 3-1 总结了 ESPN 网络的参数设置情况,包括每个网络预测分支的滤波器大小、锚框尺寸,以及锚框高宽比(高度/宽度)设置。

表 3-1 ESPN 网络中三个预测分支的参数设置

Table 3-1 Parameter configurations for three proposal branches in ESPN

网络层	Conv4_3			Conv5_3			Conv6_1		
滤波器大小/像素	3×3	5×5	7×7	3×3	5×5	7×7	3×3	5×5	7×7
锚框高度/像素	10	16	22	28	34	40	46	52	58
锚框高宽比	1:2,1:1,2:1	1:2,1:1,2:1	1:2,1:1,2:1	1:2,1:1,2:1	1:2,1:1,2:1	1:2,1:1,2:1	1:2,1:1,2:1	1:2,1:1,2:1	1:2,1:1,2:1

图 3-2a 和图 3-2b 对比了 RPN 网络与 ESPN 网络候选区域预测的过程。从图中能够看到 RPN 网络单独从最后一个卷积层预测候选框区域,而且 RPN 子网络的优化目标函数为:

$$L_{\text{RPN}} = \sum_{i=1}^N l_{\text{cls}}(p(x_i), y_i) + \lambda [y_i \geq 1] l_{\text{loc}}(b_i, \hat{b}_i) \quad (3-1)$$

其中, l_{cls} 和 l_{loc} 分别表示分类损失和定位损失, $X = \{x_i\}_{i=1}^N$ 是预测框区域的深度特征, N 是预测区域的个数, $p(x) = (p_0(x), p_1(x))$ 是预测框区域的类别概率(包括预测为背景的概率和预测为目标区域的概率), $Y = \{y_i\}_{i=1}^N$ 是预测框区域的真实类别信息, $b_i = \{b_i^x, b_i^y, b_i^w, b_i^h\}$ 是第 i 个目标真实位置信息, $\hat{b}_i = \{\hat{b}_i^x, \hat{b}_i^y, \hat{b}_i^w, \hat{b}_i^h\}$ 是第 i 个预测位置信息, λ 表示定位损失系数, $[y_i \geq 1]$ 表示当预测区域包含舰船目标时,该预测区域对定位损失函数有所贡献,否则定位损失为 0,更多 RPN 网络的损失函数定义情况见 [58]。

与 RPN 网络不同的是,ESPN 网络从多个拓展的网络子分支预测候选框区域,该子网络最终的损失函数是所有预测分支的损失函数加权和。其中,每个预测分支所对应的损失函数与 RPN 网络的损失函数形式相同。因此,ESPN 网络训练的优化目标函数可表

示为:

$$L_{\text{ESPN}} = \sum_{m=1}^M \sum_{j \in S^m} \alpha_m \cdot l_m(X_j, Y_j, B_j | W) \quad (3-2)$$

其中, W 是 ESPN 网络中需要优化的参数集合, M 是候选框提取的总通道数, S^m 是指当前预测区域属于第 m 个候选区域预测分支, X_j 、 Y_j 、 B_j 分别表示第 j 个网络分支中预测框的深度特征、类别信息、以及位置信息, l_m 是第 m 个候选框提取分支的损失函数, 其的定义形式同式 (3-1), α_m 是第 m 个候选框提取分支损失函数的计算系数, 该参数的取值表示不同尺度候选区域预测层对最终损失函数的贡献程度。

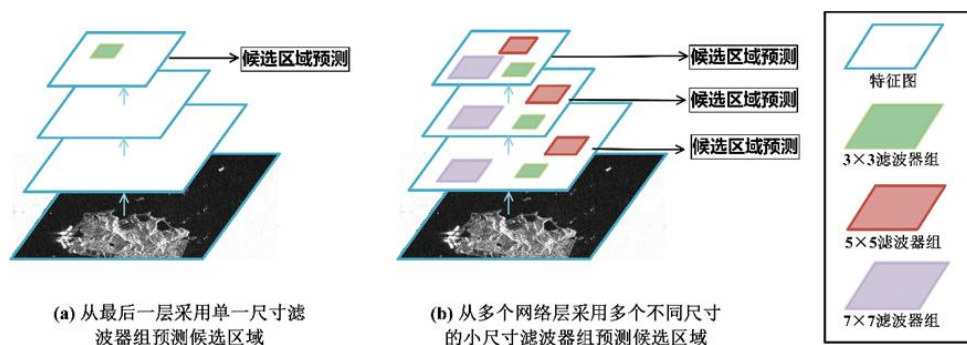


图 3-2 检测网络候选框提取机制: (a) RPN 网络; (b) ESPN 网络。

Figure 3-2 Ship proposal generation strategy in object detection: (a) RPN and (b) ESPN.

在通常情况下, ESPN 网络为了保证区域预测的查全率, 预测结果中不可避免地会存在较多虚警目标。为了提高模型的检测精度, 需要通过 ASDN 舰船目标鉴别子网络有效地剔除这些虚警区域, 保留真实舰船目标区域。在 Faster R-CNN 网络中, Fast R-CNN 鉴别网络仅通过网络结构中的最后一个卷积层输出的特征对目标进行表征。根据实际情况可知, 舰船目标检测区域中目标与周围环境具有非常强的相关性, 通常舰船目标都存在于远海海域、近海或者其他江河等大型水体中, 因此, 在对舰船目标的特征进行描述时, 目标的上下文特征也是一个不容忽视的因素, 这也是沿用了多尺度卷积神经网络中的目标鉴别子网络结构。本章目标鉴别网络 ASDN 通过融合舰船目标候选区域的深度特征及其上下文信息的方式提高目标鉴别性能。具体地, 就是对候选区域和其对应的上下文区域分别提取深度特征, 然后通过 RoIPooling 映射的方法, 将两个不同区域范围的特征输出为相同尺度的特征, 然后通过特征拼接的方式进行融合, 实现后续目标鉴别。具体的上下文特征融合实现方法如图 3-3 所示, 图中上半部分和下半部分分别展示了一个候选框示例及其所对应的上下文区域。二者通过 RoI 池化操作之后将特征图拼接, 继而合成新的融合特征, 从而提高目标鉴别子网络对目标区域的描述性能。上下文区域的选

择是在区域候选框的定位基础上，沿方位向和距离向分别延伸 1.5 倍的区域内提取其深度特征。

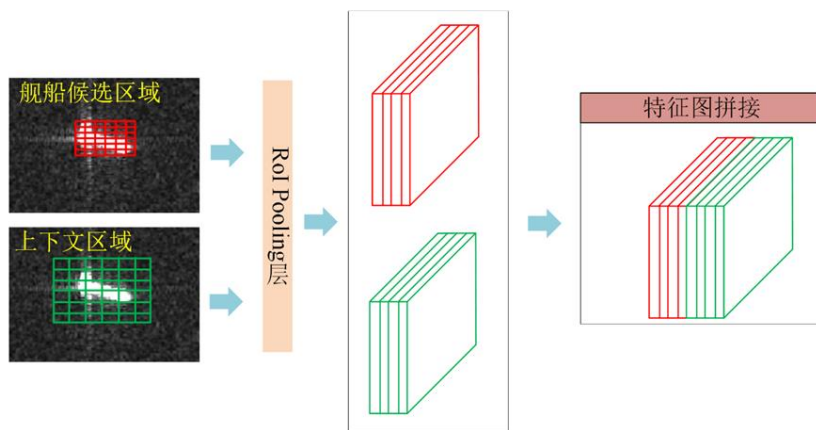


图 3-3 舰船目标候选区域深度特征与上下文特征融合

Figure 3-3 Deep Features from Ship proposals and their contextual regions are concatenated and bypassed by the RoIPooling layer

因此，基于多尺度卷积神经网络的 SAR 图像舰船目标检测过程中，网络的整体损失函数可以表示为

$$L(W, W_d) = L_{\text{ESPN}}(W) + \alpha_{M+1} L_{\text{ASDN}}(X_i, Y_i, B_i | W_d) \quad (3-3)$$

这里， α_{M+1} 是指精确目标鉴别网络 ASDN 中损失函数的权重， W_d 表示 ASDN 网络中除共享网络部分之外的参数集合。网络优化的形式可以表示为

$$W^*, W_d^* = \arg \min L(W, W_d) \quad (3-4)$$

网络训练过程中的参数更新策略采用批量随机梯度下降（Stochastic Gradient Descent, SGD）法。初始学习率设置为 0.0005，并且在训练过程中，优化函数每迭代 5000 次，学习率就降为原来的 0.1 倍。这样，经过 35,000 次迭代优化与参数更新，目标函数达到收敛状态。最后采用非极大值抑制方法剔除重复检测。

在实验过程中，本节采用参数数量与乘加运算量（Multiply-Add Computation, MAC）来衡量网络的计算成本。表 3-2 以输入 1024×768 像素图像为例，统计出了改进的多尺度卷积神经网络的结构信息、每层的参数数量，以及乘加计算量。从表中能够明显看到共享卷积层部分、候选框预测网络部分、目标鉴别网络部分的参数数量分别为 18966.2K、792.6K、57712.7K，相应的 MAC 分别为 258653.9M、4112.1M、170.4M。这表明多尺度卷积神经网络的参数和计算量主要集中在共享卷积网络层部分，对 ESPN 和 ASDN 网络层的改进几乎不以牺牲网络的计算消耗为代价。

表 3-2 改进的多尺度卷积神经网络结构、参数数量和乘加计算量消耗

Table 3-2 Detailed structure, number of parameters, and MAC when using the revised Multiscale-CNN

网络结构部分	网络层名	网络层结构	步长	输出特征图尺寸	参数数量	MAC
共享卷积层	Conv1_1	3×3 Convolution	1	$1024 \times 768 \times 64$	1.9K	1359.0M
	Conv1_2	3×3 Convolution	1	$1024 \times 768 \times 64$	41.0K	28991.0M
	Pool1	2×2 max pooling	2	$512 \times 384 \times 64$		
	Conv2_1	3×3 Convolution	1	$512 \times 384 \times 128$	81.9K	14495.5M
	Conv2_2	3×3 Convolution	1	$512 \times 384 \times 128$	163.8K	28991.0M
	Pool2	2×2 max pooling	2	$256 \times 192 \times 128$		
	Conv3_1	3×3 Convolution	1	$256 \times 192 \times 256$	327.7K	14495.5M
	Conv3_2	3×3 Convolution	1	$256 \times 192 \times 256$	655.4K	28991.0M
	Conv3_3	3×3 Convolution	1	$256 \times 192 \times 256$	655.4K	28991.0M
	Pool3	2×2 max pooling	2	$128 \times 96 \times 256$		
	Conv4_1	3×3 Convolution	1	$128 \times 96 \times 512$	1310.7K	14495.5M
	Conv4_2	3×3 Convolution	1	$128 \times 96 \times 512$	2621.4K	28991.0M
	Conv4_3	3×3 Convolution	1	$128 \times 96 \times 512$	2621.4K	28991.0M
	Pool4	2×2 max pooling	2	$64 \times 48 \times 512$		
	Conv5_1	3×3 Convolution	1	$64 \times 48 \times 512$	2621.4K	7247.8M
	Conv5_2	3×3 Convolution	1	$64 \times 48 \times 512$	2621.4K	7247.8M
	Conv5_3	3×3 Convolution	1	$64 \times 48 \times 512$	2621.4K	7247.8M
	Pool5	2×2 max pooling	2	$32 \times 24 \times 512$		
	Conv6_1	3×3 Convolution	1	$32 \times 24 \times 512$	2621.4K	18119M
共享卷积层 计算量汇总					18966.2K	258653.9M
ESPN	SPN4_3	3×3 Convolution	1	$128 \times 96 \times 6$	30.7K	339.7M
	SPN4_5	5×5 Convolution	1	$128 \times 96 \times 6$	79.9K	943.7M
	SPN4_7	7×7 Convolution	1	$128 \times 96 \times 6$	153.6K	1849.7M
	SPN5_3	3×3 Convolution	1	$64 \times 48 \times 6$	30.7K	84.9M
	SPN5_5	5×5 Convolution	1	$64 \times 48 \times 6$	79.9K	235.9M
	SPN5_7	7×7 Convolution	1	$64 \times 48 \times 6$	153.6K	462.4M
	SPN6_3	3×3 Convolution	1	$32 \times 24 \times 6$	30.7K	21.2M
	SPN6_5	5×5 Convolution	1	$32 \times 24 \times 6$	79.9K	59.0M
ESPN 网络部分 计算量汇总					792.6K	4112.1M
ASDN	RoIPooling1	7×7 RoI Pooling		$7 \times 7 \times 512$		
	RoIPooling2	7×7 RoI Pooling		$7 \times 7 \times 512$		
	RoI_concat	3×3 Convolution		$5 \times 5 \times 512$	5242.9K	118.0M
	FC	FC		4096	52428.8K	52.4M
	FC_cls	FC		2	8.2K	32.8K
	FC_bbr	FC		8	32.8K	32.8K
ASDN 网络部分 计算量汇总					57712.7K	170.4M
整个网络计算量汇总					75.66M	256.77B

3.3.3 实验结果与分析

3.3.3.1 实验准备

本实验采用加州伯克利大学贾扬青博士开发的深度学习平台 Caffe^[116]。硬件环境包括两个 32 核的 i7 处理器和一个 NVIDIA GTX-1080 GPU。

本节中的舰船目标检测采用四个常用的评价指标，分别是：精度 precision、查全率

recall、平均精度 (Average Precision, AP) 和 F_1 分数值^[117]。其中，精度的定义为

$$precision = \frac{\#(TP)}{\#(TP) + \#(FP)} \quad (3-5)$$

其中， $\#(TP)$ 和 $\#(FP)$ 分别表示正确检测与错误检测的目标个数。查全率定义为

$$recall = \frac{\#(TP)}{\#(TP) + \#(FN)} \quad (3-6)$$

这里， $\#(FN)$ 表示漏检的目标数量。AP 的定义是 PR (precision-recall curve, PRC) 曲线在查全率从 0 到 1 之间的积分。 F_1 分数定义为

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (3-7)$$

为了验证 ESPN 网络和 ASDN 网络的性能，本节的评价对比方法采用 CNN-Single-RPN 方法、CNN-Single-ROI Pooling 方法、Faster R-CNN (FRCN) 方法，以及基于 CFAR 的多阶段舰船目标检测方法 (CFAR-MS)^[118]。其中，CNN-Single-RPN 方法是用 ESPN 网络代替 Faster R-CNN 网络中的 RPN 子网络；CNN-Single-ROI Pooling 方法是用 ASDN 网络代替 Faster R-CNN 网络中的 Fast R-CNN 子网络。

3.3.3.2 在 Sentinel-1 数据上的实验结果

本实验数据来自欧空局的 C 波段 Sentinel-1 宽幅 SAR 图像，所采用的影像来自该数据集干涉宽幅 (Interferometric Wide-swath, IW) 大地距离测量 (Ground Range Detected, GRD) 模式、VH 极化方式的数据。图像沿方位向和距离向的分辨率分别为 22 米和 20 米，像元大小在两个方向均为 10 米。考虑到计算机内存限制，本节在实验过程中将原始宽幅 SAR 影像裁剪为尺寸较小的图像块。同时为了防止图像裁切边界处的舰船目标被漏检的现象发生，图像裁切过程中在边界区域预留一定的重叠区。图 3-4 展示了 SAR 影像重叠裁切的具体方式，图中红色矩形框为实际裁剪的区域，白色矩形框为划分的每个图像块大致范围，蓝色实心圆点为图像切割网格的顶点位置。为了防止舰船目标被漏检，本文在图像裁切的过程中，沿方位向和距离向均设置了 50 个像素位置的重叠度。在此基础上，实验过程中剔除没有舰船目标的图像块，并记录预留每个图像块左上角点的位置信息，便于目标检测后的图像复原。

在训练与测试数据的准备过程中，本实验数据采集自 52 幅 SAR 图像共计三万余舰船目标做为训练数据，另外八幅 SAR 图像用于模型性能测试。训练和测试数据的真值标注采用欧空局的 Sentinel 应用平台 (Sentinel Application Platform, SNAP) 以及 OpenSARShip 开发的图像标注软件插件^[41]，真值判定采用人工判读与 AIS (Automatic

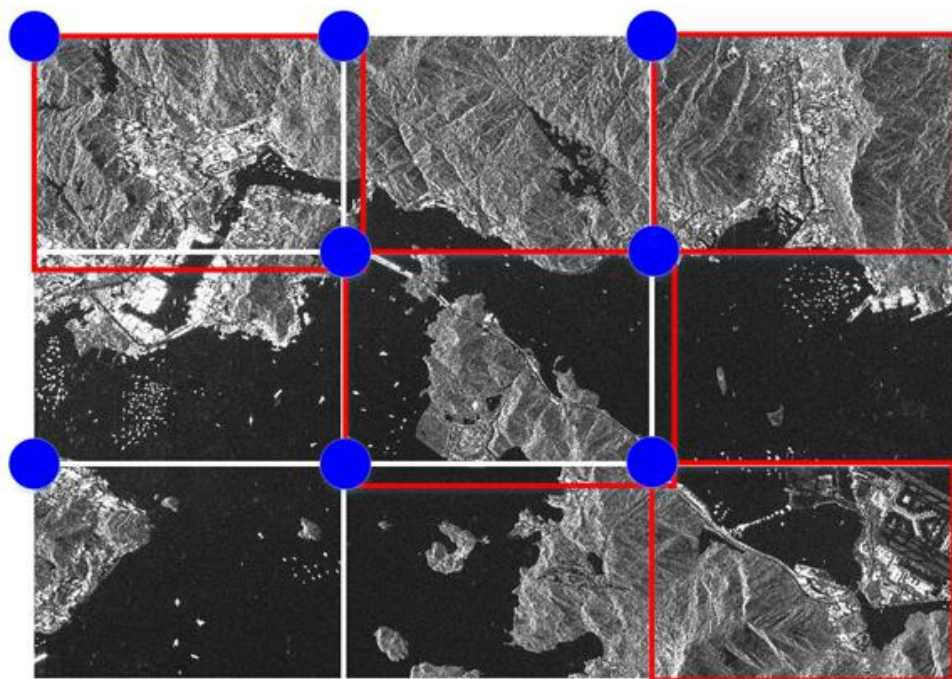


图 3-4 宽幅 SAR 图像裁剪方式

Figure 3-4 Image crop method for wide-swath SAR images

Information System) 数据^[119]相结合的方式。标注的真值以左上角与右下角图像像素坐标的形式保存。

多尺度卷积神经网络及其他四种对比方法在该数据集上的检测结果 PR 曲线如图 3-5 所示。PR 曲线下的面积表示检测的平均精度 AP 值。通过观察图中的结果，可以明显看到改进的多尺度卷积神经网络的 PR 曲线沿横轴和纵轴分别达到最长和最高的状态，因而该曲线下方所包含的面积也在所有评价方法中最大，表明本节所采用的多尺度卷积神经网络方法能够取得最大的 AP 值。同时，也不难发现其他基于卷积神经网络的方法所取得的 AP 值也高于 CFAR-MS 方法。

表 3-3 列出了 Sentinel-1 数据集在 $\text{IoU}=0.3$ 时的量化检测性能指标。从表中所得出的结果可以看到：

- 改进的多尺度卷积网络的检测性能明显优于其他基于卷积神经网络的目标检测方法和传统基于 CFAR 的多阶段目标检测方法的检测性能；
- 改进后的 CNN-Single-ROI Pooling 方法和 CNN-Single-RPN 方法检测性能优于 Faster R-CNN 方法；
- 基于卷积神经网络的舰船目标检测方法性能优于 CFAR-MS 多阶段目标检测方

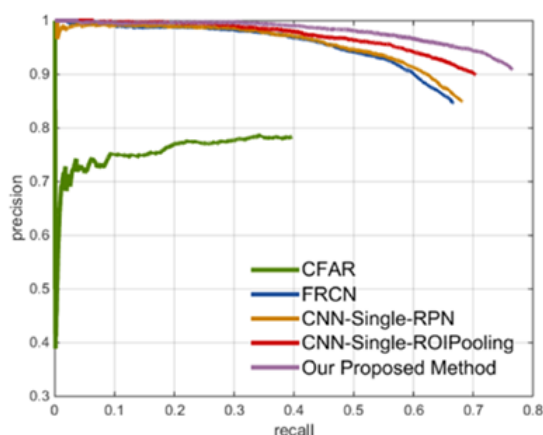


图 3-5 在 Sentinel-1 实测 SAR 图像上的 PR 曲线

Figure 3-5 PR curves over Sentinel-1 real SAR data

法；

- 基于卷积神经网络的检测方法在检测效率上比 CFAR-MS 多阶段目标检测方法更高。

表 3-3 在 Sentinel-1 实测数据上的量化结果

Table 3-3 Performance comparison of different methods for the Sentinel-1 real data

方法	查全率	精度	平均精度	F_1 值	每幅图平均耗时 (s)
CFAR-MS	0.3977	0.7830	0.3123	0.5275	2550
FRCN	0.6669	0.8432	0.5812	0.7447	105
CNN-Single-RPN	0.7107	0.8964	0.6519	0.7928	113
CNN-Single-ROI Pooling	0.6833	0.8498	0.6069	0.7575	108
改进的多尺度卷积神经网络	0.7719	0.9022	0.7151	0.8320	115

除了上述量化指标评价之外,为了更加直观地展现所采用的改进多尺度卷积神经网络在 Sentinel-1 舰船检测数据集上的优越性,本节通过图 3-6 分别展示了改进多尺度卷积神经网络方法和 CFAR-MS 多阶段舰船目标检测方法的检测结果。图中蓝色、绿色和红色的矩形框分别表示舰船目标真值标注、正确检测目标和虚警目标。为了更加清晰地显示真实的检测效果,图中的黄色矩形框所在区域的检测结果被放大化。图 3-6a 为本节所采用的多尺度卷积神经网络的检测效果图,图 3-6b 为 CFAR-MS 多阶段舰船检测方法的检测效果图,能够看到图中所展示区域为近港区域而且目标密集排布。图 3-7 展示了远海海域的检测视觉效果图,其中,图 3-7(a-1)~(a-4) 为改进多尺度卷积网络的检

测结果，图 3-7(b-1)~(b-4) 为 CFAR-MS 方法的检测结果。从两幅图的对比效果中可以明显看到，无论是近港区域中的密集排布目标，还是在远海海域中的目标，改进多尺度卷积神经网络的检测性能都明显优于 CFAR-MS 多阶段舰船目标检测方法。

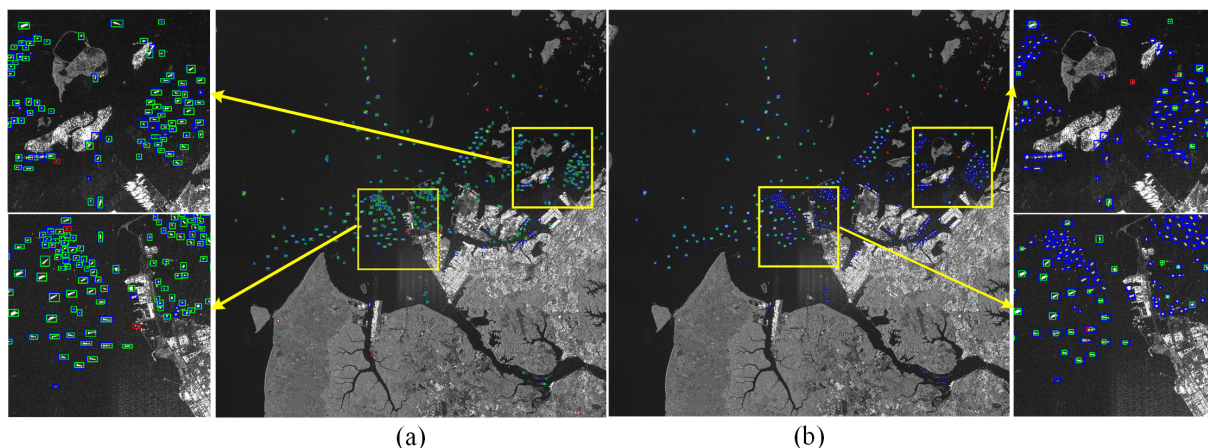


图 3-6 在 Sentinel-1 影像上新加坡马六甲海峡区域的可视化检测结果，其中 (a) 改进的多尺度网络；(b) CFAR-MS。在两幅子图中，被放大的黄色矩形框区域能够更清晰地显示检测结果，绿色、红色和蓝色框分别指示正确检测的目标、错误检测的目标和标注真值。

Figure 3-6 Ship detection results with (a) Revised Multiscale-CNN and (b) CFAR-MS for an image block cropped from the wide-swath Sentinel-1 SAR imagery over the Strait of Malacca, Singapore. The green box indicates the correctly detected targets, the red indicates false alarms, and the blue represents the ground-truth.

3.3.3.3 在高分三号数据上的处理结果

对于高分三号数据集，本节共标注了 20 幅 NSC 成像模式、极化方式为 VH 的影像，像元大小沿方位和距离向分别为 20 米和 5 米。模型训练数据来自 16 幅高分三号影像。改进多尺度卷积神经网络及其他四种对比方法在该数据集上的 PR 曲线如图 3-8 所示。PR 曲线沿横轴的积分表示检测结果的 AP 值。从两幅结果图中可以明显看到改进多尺度卷积神经网络的 PR 曲线沿横轴和纵轴分别能够达到最长和最高，因此该曲线下包含的面积也最大，表明该方法取得的 AP 值是所有方法中最大的。

表 3-4 列出了高分三号数据集在 $\text{IoU}=0.3$ 时的量化检测性能指标。通过观察表 3-4，可以看到：

- 改进多尺度卷积神经网络的方法的检测性能明显优于其他方法的检测性能；
- CNN-Single-ROI Pooling 方法和 CNN-Single-RPN 方法检测性能优于 Faster R-CNN 方法；

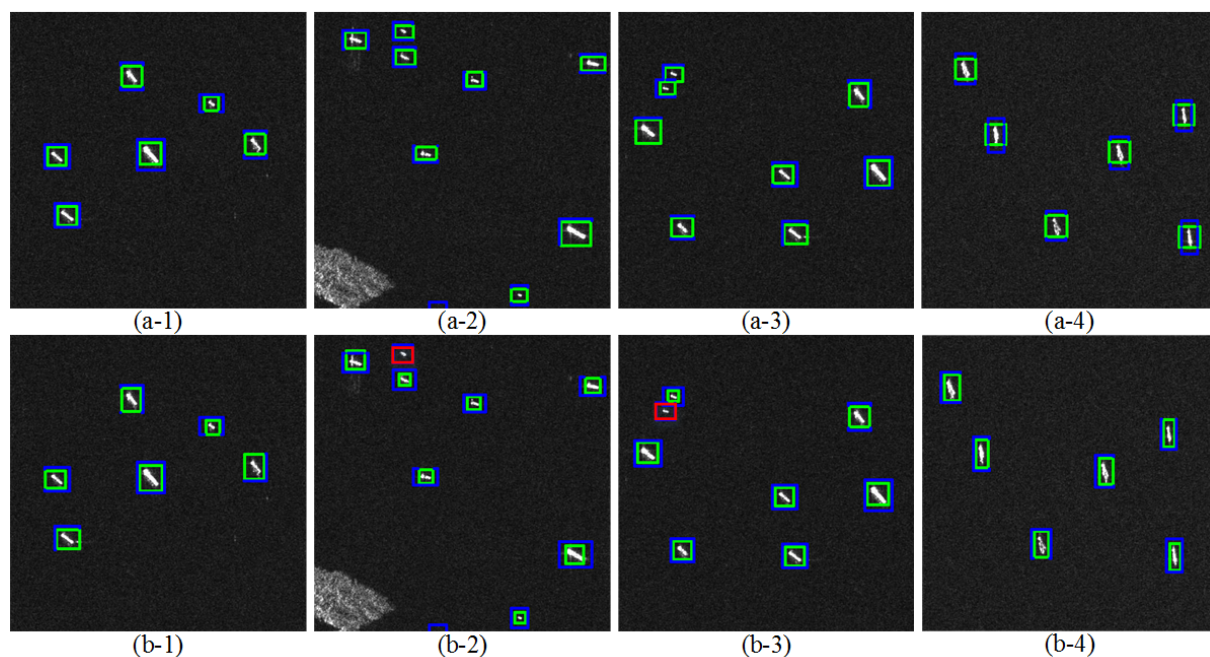


图 3-7 在 Sentinel-1 影像上远海海域的 SAR 图像舰船目标检测结果。(a-1)-(a-4) 展示了改进多尺度卷积神经网络的检测结果；(b-1)-(b-4) 为 CFAR-MS 方法的检测结果。绿色、红色和蓝色框分别表示正确检测的目标、错误检测的目标和标注真值。

Figure 3-7 Ship detection results in offshore areas of Sentinel-1 images. (a-1)-(a-4) exhibit the visualization results by using the proposed Coupled-CNN method. (b-1)-(b-4) show the detection results of the CFAR-MS method. The green box indicates the correctly detected targets, the red indicates false alarms, and the blue represents the ground-truth.

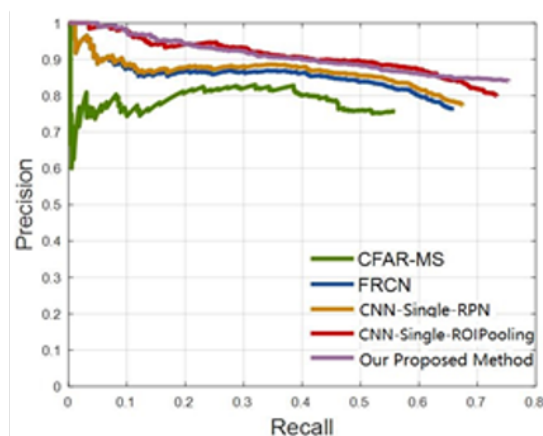


图 3-8 在高分三号实测数据上的 PR 曲线

Figure 3-8 PR curve over GF-3 real data

- 基于卷积神经网络的目标检测方法的检测性能优于 CFAR-MS 舰船目标检测方法；
- 在检测效率上，基于卷积神经网络的目标检测方法性能优于基于 CFAR-MS 的方法。

表 3-4 在高分三号实测数据上的量化检测结果

Table 3-4 Performance comparison of different methods for the GF-3 data set

方法	查全率	精度	平均精度	F1 值	每幅图平均耗时 (s)
CFAR-MS	0.5582	0.7562	0.4832	0.6423	1630
FRCN	0.6710	0.7597	0.5772	0.7126	85
CNN-Single-RPN	0.7433	0.7906	6784	0.7662	86
CNN-Single-ROIPooling	0.6887	0.7687	0.5997	0.7265	89
改进的多尺度卷积神经网络	0.7719	0.9022	0.7151	0.8320	115

为了更加直观地展现多尺度卷积神经网络方法在高分三号舰船检测数据集上的优越性，本节采用图 3-9 展示了该方法和 CFAR-MS 检测方法的检测结果。为了更加清晰地显示两种方法的检测效果，图中黄色矩形框所在区域的检测结果被放大。图 3-9a 为改进多尺度卷积神经网络的检测效果图，图 3-9b 为 CFAR-MS 检测方法的检测效果图。值得一提的是，该图所展示的区域大多为近岸区域而且大多由密集排布的小目标构成。图 3-10 展示了两种方法在远海海域中的舰船目标检测效果，其中，图 3-10(a-1)~(a-4) 显示了改进多尺度卷积网络的检测结果，图 3-10(b-1)~(b-4) 显示了 CFAR-MS 检测方法所取得的视觉效果图。从两幅图的对比效果中明显可以看到，无论是近岸区域中的密集排布目标，还是在远海海域中的舰船目标，改进多尺度卷积神经网络的检测性能都明显优于 CFAR-MS 舰船目标检测方法。

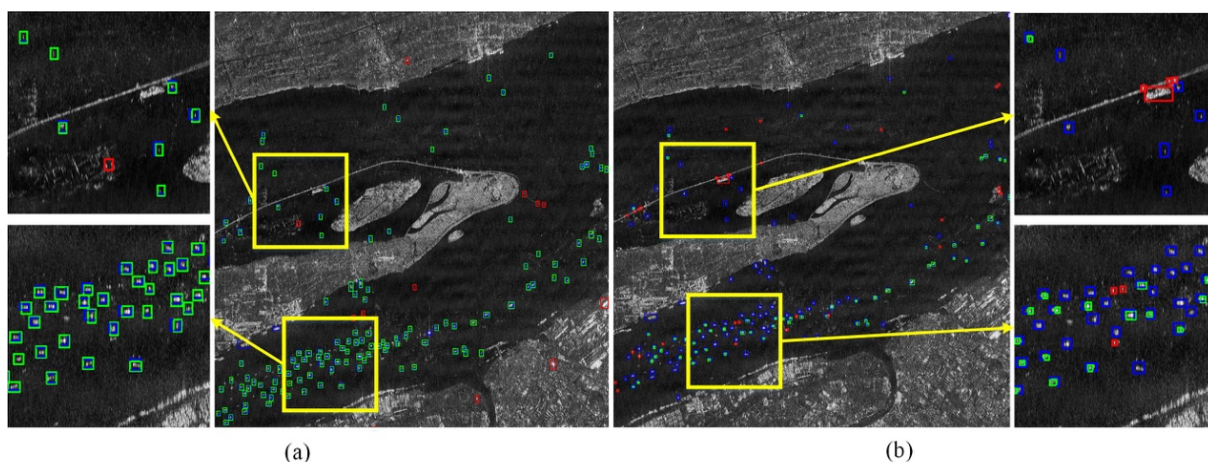


图 3-9 高分三号实测数据上的可视化舰船目标检测结果，其中，(a) Revised Multiscale-CNN；(b) CFAR-MS。绿色、红色和蓝色框分别表示正确检测的目标、错误检测的目标和标注真值。

Figure 3-9 Ship detection results with (a) Revised Multiscale-CNN and (b) CFAR-MS from a GF-3 SAR image. The green box indicates the correctly detected targets, the red indicates false alarms, and the blue represents the ground-truth.

3.4 结合级联网络与脉冲余弦变换的舰船目标精细化提取

3.4.1 研究动机

虽然上述多尺度卷积神经网络目标检测方法在 SAR 图像舰船目标检测中取得了检测性能上的明显提升，但是最终回归所得到的矩形检测框精确度仍然不足，在目标密集区域更是存在大量检测框交叉、漏检目标较严重的问题。这也就是说相对目标来说，检测所得的矩形框比较粗糙，检测框与目标的最小外接矩形逼近程度不高。同时，由于在靠港/岸区域等一些舰船目标分布密集区域的检测框交叉问题，非极大值抑制操作会导致在这些区域中出现较严重的漏检现象。此外，在陆地区域，由于某些陆地目标，如小型建筑物，其几何形状、形态学特征等与海上舰船目标在中低分辨率 SAR 图像中的相似性较高，也会被检测器误认为是舰船目标，导致检测结果中存在部分陆地虚警。究其原因，这一方面是由于人工标注误差所带来的网络学习偏差导致；另一方面是由于舰船目标的密集排布对检测框造成的干扰较强。通过改进的多尺度卷积神经网络所取得的 SAR 图像舰船目标检测结果中所存在的问题展示在图 3-11 中。其中，图 3-11a 展示了在舰船目标密集排布区域中多尺度卷积神经网络方法所取得的检测结果，从图中可以明显看到，由于目标与目标之间的距离较近，而检测框又比较粗糙，导致相邻两个检测框之间重合度较高，与此同时，漏检现象也较为严重。图 3-11b 展示了陆地虚警现象，这是由于陆地上某些地物目标的几何形状或者形态学特征等在中低分辨率 SAR 图像中与

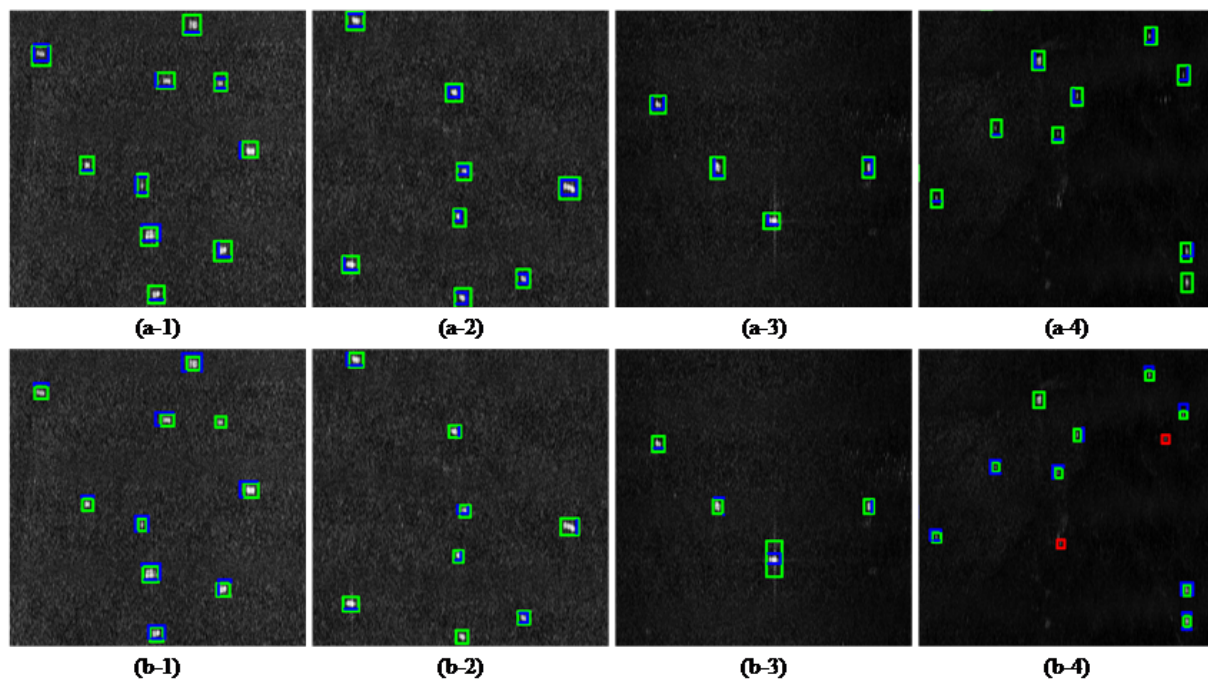


图 3-10 在远海海域高分三号实测数据上的 SAR 图像舰船目标检测结果。(a-1)-(a-4) 展示了多尺度卷积神经网络的检测结果；(b-1)-(b-4) 为 CFAR-MS 方法的检测结果。绿色、红色和蓝色框分别表示正确检测目标、错误检测目标和标注真值。

Figure 3-10 Ship detection results in offshore areas of GF-3 images. (a-1)-(a-4) exhibit the visualization results by using Cascade-CNN. (b-1)-(b-4) show the detection results of CFAR-MS. The green box indicates the correctly detected targets, the red indicates false alarms, and the blue represents the ground-truth.

舰船目标相似度较高所引起的。

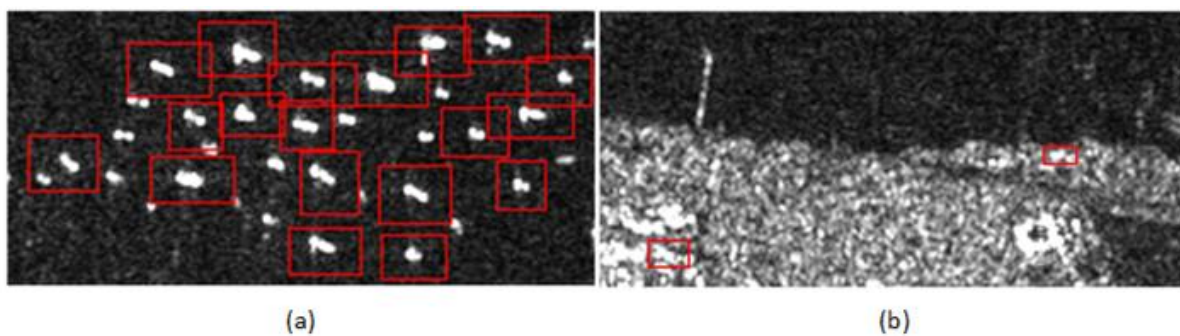


图 3-11 改进多尺度卷积神经网络方法检测结果中存在的问题分析

Figure 3-11 Demerit analysis from SAR ship detection results based on revised Multiscale-CNN

虽然与传统的多阶段 SAR 舰船检测方法相比，多尺度卷积神经网络在尺寸较小且

目标密集排布的近港/岸区域中取得了较好的检测效果。但是该网络模型仍然存在以下缺陷：

- 从信息挖掘的角度来说，基于多尺度卷积神经网络的方法仅仅利用了图像的时域信息，而忽略了图像的频域信息，导致其对图像的表征具有一定的局限性；
- 多尺度卷积神经网络方法对图像的空间域信息挖掘还不够充分；
- 蕴含在影像数据中的其他源的信息没有得到充分有效的利用。

考虑到实际情况中，舰船目标主要分布在水体环境中，目标会随着水体的浮动而有一定程度的抖动，并非完全静止。由于频域信息对动目标具有较强的表征性能，SAR 舰船目标的频域信息是一个不可忽视的因素。为了弥补多尺度卷积神经网络在 SAR 图像舰船目标检测过程中忽视频域信息的这一缺陷，本章采用级联网络结合基于脉冲余弦变换（Pulse Cosine Transformation, PCT）的视觉注意机制模型用于 SAR 图像中舰船目标的精细化提取。与此同时，为了更加充分地挖掘图像的空间域信息，本节采用深度学习中的级联结构^[120]和多尺度策略结合的方法提取舰船目标候选区域，将所提取到的疑似目标区域作为初始粗略检测结果。在此基础上，通过在一个局部自适应区域中采用基于 PCT 的频域处理方法，并结合数字高程模型（Digital Elevation Model, DEM）数据剔除陆地虚警，实现 SAR 图像中舰船目标的精细化提取。

本节通过结合级联卷积神经网络和基于脉冲余弦变换的视觉关注方法实现 SAR 图像中舰船目标的精细化提取。该方法的两个主要构成部分包括：级联卷积神经网络用于 SAR 图像中舰船目标粗检测；基于脉冲余弦变换的视觉关注方法用于对疑似目标区域实施精提取，从而实现 SAR 图像中舰船目标的精细化提取。在该方法中，为了有效利用图像的空间域信息，级联卷积神经网络以多尺度网络作为主干网络，提取 SAR 图像中的疑似舰船目标区域。该方法通过多次重复基础网络中的目标鉴别子网络，逐步提高检测结果的准确度。基于脉冲余弦变换的视觉关注方法对级联卷积神经网络所提取的疑似候选框通过频域滤波等一系列图像处理操作，使得 SAR 图像中的舰船目标得以凸显，从而达到优化检测结果的目的。已有的研究表明，视觉显著性产生于人脑视觉皮层 V1 区初级视觉皮层中同类神经元的侧抑制过程，视觉注意机制方法能够模仿自底向上的人类视觉注意机制。与此同时，由于该方法的频域处理特性，从而有利于靠/近岸舰船目标的筛选与精确定位。为了减少漏检以及减少模型计算量，基于 PCT 的目标精准化提取在一个局部自适应区域中进行处理，该区域在疑似目标的局部范围基础上自适应扩大视觉注意机制处理在图像中的区域范围。该方法的总体思路如图 3-12 所示，本节将级联卷积神经网络的检测结果作为疑似目标区域，然后在一个局部自适应区域内，通过基于 PCT 的频域视觉注意模型进行舰船目标精准化提取。从整体上看，该方法同时结合了 SAR 图像的时域信息和频域信息。

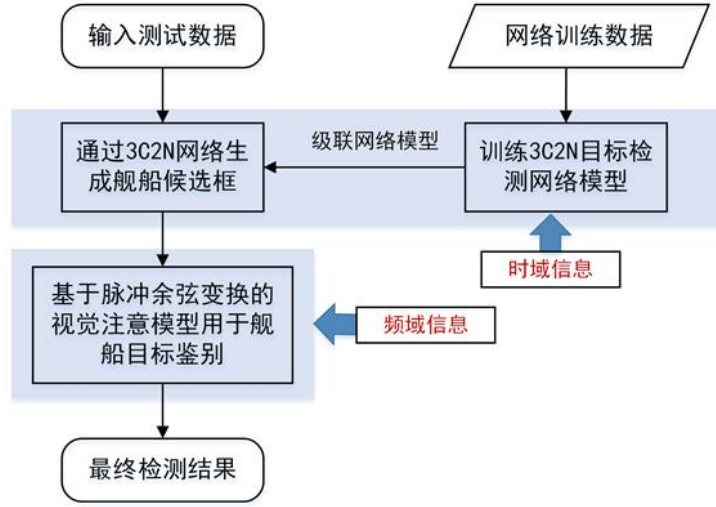


图 3-12 结合级联网络与脉冲余弦变换方法用于 SAR 图像舰船目标检测的整体思路

Figure 3-12 Ideas on the proposed cascade network and PCT method for SAR ship detection

3.4.2 方法详介

3.4.2.1 级联卷积神经网络模型

为了提高目标检测效果，本节所采用的级联卷积神经网络（3C2N）由一个 ESPN 网络和两个 ASDN 网络（ASDN-1 和 ASDN-2）通过级联结构组成，并且三个网络共享卷积特征层用于图像特征学习。图 3-13 所示为本节级联卷积神经网络的网络结构图。图中 H 和 W 分别为训练图像的高度和宽度。ESPN 网络、ASDN-1 网络，以及 ASDN-2 网络的结构沿用了上节所述多尺度卷积神经网络中的对应网络结构。其中，ESPN 网络中目标候选框提取从多个网络层，并在每个预测网络层中通过多个不同尺度的小卷积核组完成候选框预测。

级联卷积神经网络的训练是在一个一体化的网络框架中完成，网络训练过程中所采用的损失函数为各子网络损失函数的加权和。因此，本节中级联卷积神经网络所采用的优化目标函数在数学上可以表示为

$$\begin{aligned}
 L(\Theta_p, \Theta_{d1}, \Theta_{d2}) = & L_{\text{ESPN}}(\Theta_p) + \alpha_{M+1} \cdot \sum_{t \in S^{M+1}} l_{\text{ASDN-1}}(X_t, Y_t, B_t | \Theta_{d1}) \\
 & + \alpha_{M+2} \cdot \sum_{t \in S^{M+2}} l_{\text{ASDN-2}}(X_t, Y_t, B_t | \Theta_{d2})
 \end{aligned} \quad (3-8)$$

其中， $L_{\text{ESPN}}(\Theta_p)$ 和 Θ_p 分别为 ESPN 子网络的损失函数和网络参数集合， α_{M+1} 和 α_{M+2} 分别是 ASDN-1 网络和 ASDN-2 网络的损失函数 $l_{\text{ASDN-1}}$ 和 $l_{\text{ASDN-2}}$ 的系数， Θ_{d1} 和 Θ_{d2} 分别是 ASDN-1 网络和 ASDN-2 网络除共享卷积层之外的参数集合。

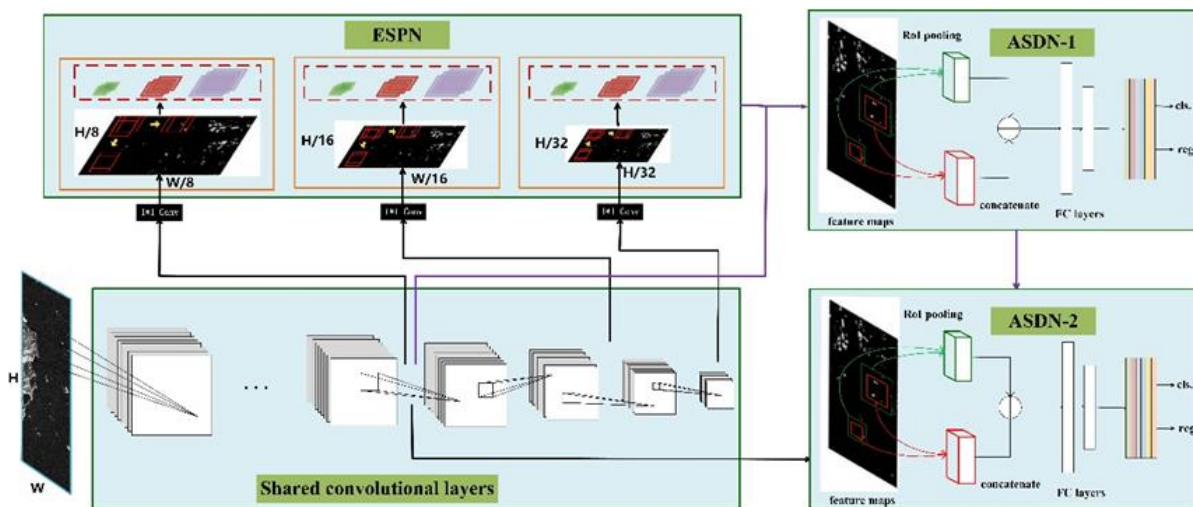


图 3-13 级联卷积神经网络的网络结构图，共包含一个候选区域提取子网络和两个目标鉴别子网络，且三个子网络共享卷积网络特征提取层。图中“FC”指代全连接层，“cls.”和“reg.”分别表示分类（包含与不包含舰船目标的区域）损失和回归损失。

Figure 3-13 Architecture of the cascade network, which consists of one ship proposal network and two cascade ship discrimination networks. All of them share some convolutional layers for feature learning. In this figure, “FC” represents fully connected layer, “cls.” and “reg.” denote the classification (ship-region or non-ship region) and bounding box regression, respectively.

图 3-14 比较直观地展示了深度卷积神经网络模型提取舰船目标候选区域过程的具体实施细节。图中上面一行依次展示了训练数据与舰船目标真值获取（舰船目标的位置信息），以及训练数据扩充方法，通过级联卷积神经网络训练所得的模型用于从测试 SAR 影像中获取疑似舰船目标位置。然后，测试数据首先通过图像重叠裁切的方法被切割成一个个小尺寸图像块，再将这裁切之后的图像块输入到网络模型中进行测试，最后将来自同一幅影像图像块的检测结果通过图像拼接的方式得到整幅大图的检测结果，该过程展示在图 3-14 中的下面一行。

3.4.2.2 脉冲余弦变换方法

脉冲余弦变换 (PCT) 模型通过白化 (归一化) 离散余弦变换 (Discrete Cosine Transform, DCT) 系数的幅度值来抑制视觉空间中出现次数较多的视觉特征，从而加强出现次数较少的视觉特征。脉冲余弦变换方法的思路试图使得模型能够尽可能精确地模拟大脑中 V1 区同类视觉神经元之间的侧抑制作用。因此，该模型能够高效地计算输入 SAR 图像场景中的视觉显著点或者视觉显著区域^[121, 122]。模型通过频率域中信号的抑制增强变换，使得在逆变换得到的时域图像中，舰船目标得以凸显。以下是针对本章的问题在

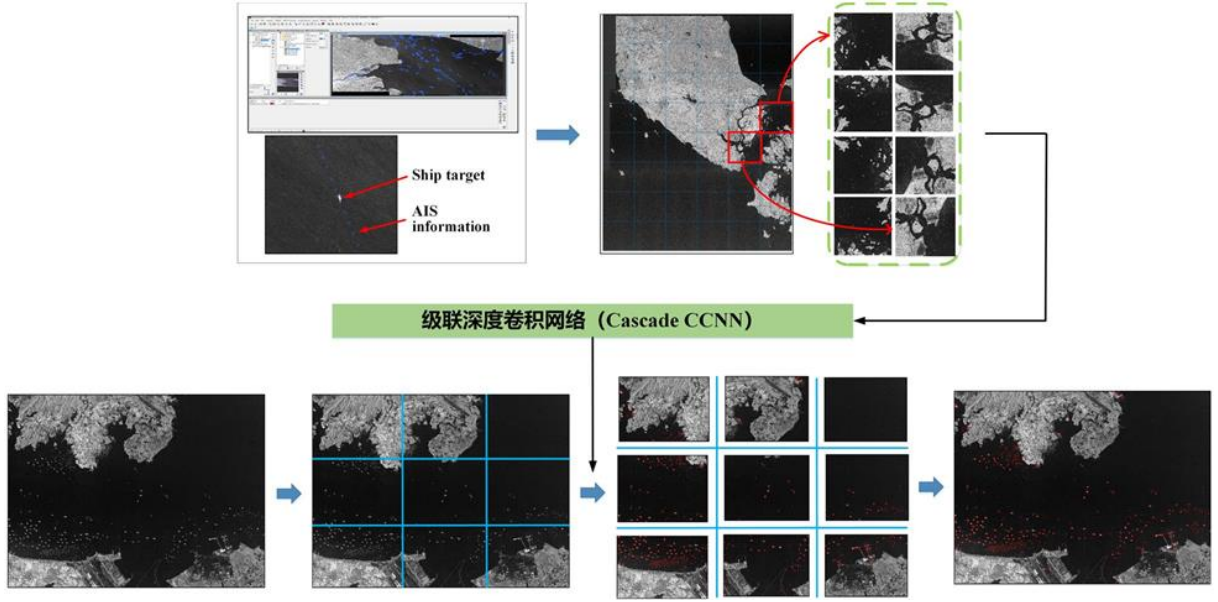


图 3-14 级联网络模型提取粗略舰船目标候选框

Figure 3-14 Ship proposals are generated by using the cascade network

具体应用时的算法流程。

给定通过自适应变换得到的目标局部扩展区域的 SAR 图像 $I \in \mathbb{R}^{5h \times 5w}$ (这里 h 和 w 分别表示疑似舰船目标区域的高度和宽度), 首先通过下式对图像进行平滑处理:

$$I_{fl}(x, y) = \begin{cases} I(x, y), & \text{if } I(x, y) \geq T_1 \\ T_1, & \text{otherwise} \end{cases} \quad (3-9)$$

其中, 阈值 T_1 由下式确定

$$T_1 = \mu(I_{fl}) + \alpha \cdot \sigma(I_{fl}) \quad (3-10)$$

这里, $\mu(I_{fl})$ 表示局部自适应区域内的图像 I_{fl} 的均值, $\sigma(I_{fl})$ 表示图像 I_{fl} 的方差, $\alpha = 0.6$ 是一个自定义的经验常数。

然后, 通过二维离散余弦变换 (Discrete Cosine Transform, DCT) 获得图像的频域特性表征, 具体的图像变换方式为

$$\begin{aligned} I_c(u, v) &= \mathcal{C}(I_{fl}(x, y)) \\ &= a_u a_v \sum_{x=0}^{5h-1} \sum_{y=0}^{5w-1} I_{fl}(x, y) \cdot \cos \frac{(2x+1)u\pi}{2 \cdot 5h} \cdot \cos \frac{(2y+1)v\pi}{2 \cdot 5w} \end{aligned} \quad (3-11)$$

其中, $\mathcal{C}(\cdot)$ 表示二维 DCT 变换操作, 且

$$a_u = \begin{cases} \frac{1}{\sqrt{5h}}, & \text{if } u = 0 \\ \sqrt{\frac{2}{5h}}, & 1 \leq u \leq 5h - 1 \end{cases} \quad (3-12)$$

$$a_v = \begin{cases} \frac{1}{\sqrt{5w}}, & \text{if } v = 0 \\ \sqrt{\frac{2}{5w}}, & 1 \leq v \leq 5w - 1 \end{cases} \quad (3-13)$$

这里, $0 \leq x \leq 5h - 1$, $0 \leq y \leq 5w - 1$, $0 \leq u \leq 5h - 1$, 并且 $0 \leq v \leq 5w - 1$ 。

然后, 经过余弦变换的图像将在频率域被量化为 1, 0 和 -1。这个操作模仿了自然界中生物的视觉机制作用, 1 和 -1 变换分别表示对相应位置处信号的激活和反激活, 0 变换表示抑制该信息。图像在频率域量化的过程可以用数学公式表达为

$$I_{sn}(u, v) = \text{sign}(I_c(u, v)) = \begin{cases} -1, & \text{if } I_c(u, v) < 0 \\ 0, & \text{if } I_c(u, v) = 0 \\ 1, & \text{if } I_c(u, v) > 0 \end{cases} \quad (3-14)$$

这里, $\text{sign}(\cdot)$ 表示符号函数。图像 I_c 中正的 DCT 系数被量化为 +1, 负的 DCT 系数被量化为 -1, DCT 系数为 0 时保留原始数据。经过量化处理之后的图像经过二维反离散余弦变换 (Inverse Discrete Cosine Transform, IDCT), 使得图像反变换到空间域, 具体变换方式为

$$\begin{aligned} I_{inv}(x, y) &= \text{abs}(\mathcal{C}^{-1}(I_{sn}(u, v))) \\ &= \text{abs}\left(\sum_{u=0}^{5h-1} \sum_{v=0}^{5w-1} a_u a_v I_{sn}(u, v) \cdot \cos \frac{(2x+1)u\pi}{2 \cdot 5h} \cdot \cos \frac{(2y+1)v\pi}{2 \cdot 5w}\right) \end{aligned} \quad (3-15)$$

这里,

$$\text{abs}(\eta) = \begin{cases} \eta, & \text{if } \eta \geq 0 \\ -\eta, & \text{otherwise} \end{cases} \quad (3-16)$$

再通过图像高斯滤波得到视觉显著图, 高斯滤波操作可以表示为

$$I_{saliency} = F_g \otimes [I_{inv}]^2 \quad (3-17)$$

其中, F_g 表示二维低通高斯滤波器, \otimes 表示二维卷积运算。最后, 对滤波后的图像进行二值化处理, 其数学表示为

$$I_{bi}(x, y) = \begin{cases} 1, & \text{if } I_{saliency} \geq T_2 \\ 0, & \text{otherwise} \end{cases} \quad (3-18)$$

其中, 阈值 $T_2 = \mu(I_{saliency}) + \beta \cdot \sigma(I_{saliency})$, $\mu(I_{saliency})$ 和 $\sigma(I_{saliency})$ 分别是显著性图像 $I_{saliency}$ 中所有像素值的平均值和标准差, $\beta = 2.5$ 为经验常数。

脉冲余弦变换模型模仿自底向上的生物视觉注意形成机制, 文献 [123, 124] 的作者已经验证了该方法在 SAR 图像解译中的有效性。在生物学中, 初级视觉皮层 V1 能够天然地产生自底向上的视觉显著图。对于其中响应最强烈的那些神经元来说, 它们的感受野对应的场景区域成为视觉注意焦点的可能性也是最大的。该观点认为, 视觉显著性的形成是同类神经元侧抑制作用的结果, 也就是说一个脉冲发放的神经元会抑制其周围同类神经元信号的发放。如果一个与周围明显不同的视觉特征被一个神经元检测到, 因它没有受到周围同类神经元的抑制, 所以发放率较高。与之相反, 如果一个与周围相同的视觉特征被一个神经元检测到, 由于它受到周围同类神经元的抑制, 其脉冲发放率也会大大降低。因此, 高发放率的神经元总是会出现在视觉特征较为突出的位置。与二维傅里叶变换类似, 二维离散余弦变换通过不同频率和不同方向的周期信号 (基向量) 来表征图像, 因此, 图像的二维离散余弦变换系数中蕴含同类视觉特征在空间上出现的统计信息, 数值较大的二维离散余弦变换系数表明相对应的频率和方向上的视觉特征在空间上出现的频次较多。

3.4.2.3 基于脉冲余弦变换的目标精细化提取

本节在级联卷积神经网络提取疑似舰船目标区域的基础上, 采用基于脉冲余弦变换的频域视觉注意模型实现 SAR 图像中舰船目标的精准化提取。图 3-15 展示了具体的算法实现流程图。

基于脉冲余弦变换的目标精细化提取具体实施细节如下文所述。

1. 提取初始舰船目标候选区域。这里的候选区域是指从级联卷积神经网络的检测结果中所提取到的粗略检测结果 $[x_s, y_s, h, w]$, 其中 x_s 和 y_s 分别表示初始舰船目标候选矩形框区域的左上角角点在方位向和距离向上的像素坐标, h 和 w 分别表示所对应的候选矩形框区域的高度和宽度, 也就是候选矩形区域在方位向和距离向上所占据的像素个数。
2. 局部自适应扩展候选区。由级联卷积神经网络的检测结果可以看到, 漏检的目标大多数都在检测结果的长度和宽度方向分别延伸 2 倍的长度范围内。因此, 局部

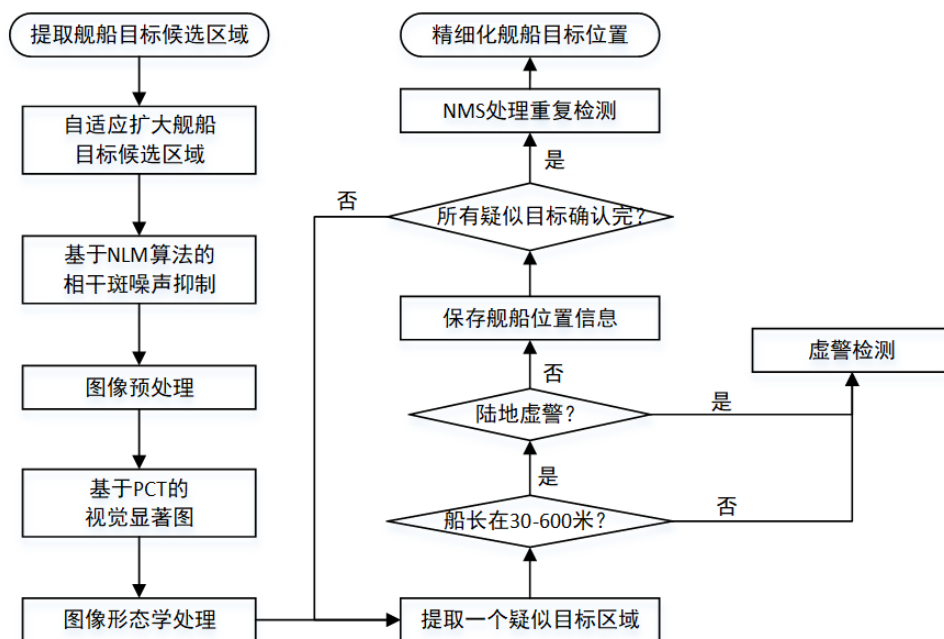


图 3-15 基于脉冲余弦变换的目标精细化提取流程图

Figure 3-15 Workflow of the PCT-based visual attention for accurate ship extraction

自适应扩展区域定义在以初始候选区域中心点为中心,上下左右四周分别延伸至相应候选框长度和宽度 2 倍的区域。从而得到高度和宽度分别为 $5h$ 和 $5w$ 的局部自适应扩展区域,该区域在图像中的坐标位置表示为 $[x_s-2h+1, y_s-2w+1, 5h, 5w]$ 。特别地,如果扩展框的边界超出了图像尺寸范围,则以原始图像边界作为扩展区域的对应边界。图 3-16 展示了局部自适应扩展区域,红色矩形框表示步骤 1 中提取到的初始舰船目标候选区域,绿色矩形框表示局部自适应扩展之后的矩形区域。

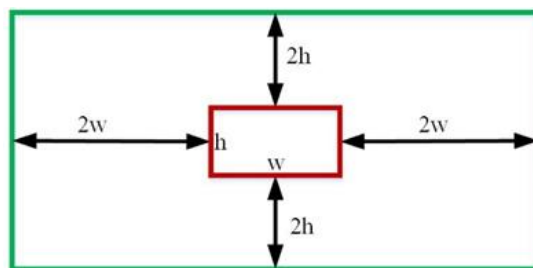


图 3-16 局部自适应候选区域设计方法

Figure 3-16 Designation of the regional adaptive proposal regions

3. 在图像中的局部自适应候选区域内进行相干斑噪声抑制。本节采用非局部均值

(Non-local Mean, NLM) 算法来抑制 SAR 图像中的相干斑噪声, 该算法是常用的有效抑制 SAR 图像相干斑噪声的算法之一^[125]。该操作在数学上通过下式表达

$$I_{NL} = NonLocalMean(R_{ad}) \quad (3-19)$$

其中, $NonLocalMean(\cdot)$ 表示非局部均值算法, I_{NL} 为相干斑噪声抑制后的局部自适应扩展区域。

4. 图像预处理。图像预处理的目的是平滑海杂波和陆地噪声的影响, 并使研究对象得以加强。为了削弱部分陆地区域的散射以及海杂波, 并使得舰船目标部分的像素值得到加强, 本节在步骤 3 的基础上, 对相干斑噪声抑制后的局部自适应矩形区域 I_{NL} 逐像素进行下列变换。对于像素 $\forall p \in I_{NL}$, 则其像素值的变换方式为

$$f(p) = \frac{[\mu_{local}]^2}{2 \times [\sigma_{ad}]^2} \quad (3-20)$$

这里, μ_{local} 为以当前像素为中心的 3×3 像素数大小区域内的均值, σ_{ad} 为局部自适应扩展矩形区域内所有像素值的标准差。这样经过处理之后得到的图像表示为 I_{pp} 。

5. 基于 PCT 的视觉注意变换。通过在频域中的处理, 使得目标的显著性凸显是该步骤的目的。值得一提的是该步骤处理之后, 在一个局部自适应扩展区域内可能会存在多艘舰船目标, 本节将舰船所在像素值设置为 1, 其余设置为 0, 即得到二值图像 I_{bi} 。具体处理方式见 3.4.2.2 部分。
6. 图像的形态学处理。首先确定连通区域, 然后计算每个连通区域的面积。当某一个连通区域的面积小于 10 个像素时, 该连通区域被认为是非舰船目标所在区域。
7. 检测结果校正。为了进一步提高检测的准确性, 本节首先考虑将舰船尺寸作为约束条件之一。在具体的实施过程中, 将舰船高度和宽度小于 30 米或者大于 600 米的目标认为是非舰船目标。其次, 为了降低陆地虚警所带来的影响, 本节借助 SRTM DEM 数据, 对于检测框位置落在陆地区域内的目标认为是非舰船目标。
8. 提取优化后的检测结果。记录上述步骤中提取到的检测结果在其局部自适应扩展区域内的坐标, 并将其表示为 $[x_{rb}, y_{rb}, h_r, w_r]$ 。
9. 坐标转换。由于前面所提取到的目标检测框是以自适应矩形检测区域作为参考的位置坐标, 因此, 本节通过位置变换得到其在整个大图中的像素坐标, 坐标转换公式为

$$[x_{rs}, y_{rs}, h_r, w_r] = [x_{rb}, y_{rb}, h_r, w_r] + [x_s - 2h, y_s - 2w, 0, 0] \quad (3-21)$$

10. 返回步骤 8 直至自适应矩形扩展区域内的所有目标都优化完毕。
11. 返回步骤 1 处理另一个由级联卷积神经网络提取到的疑似目标区域。
12. 通过非极大值抑制的方法 (NMS 算法) 去除重复检测。

算法 3-1 描述了舰船目标精细化提取过程中的细节信息, 算法中 PS_{az} 和 PS_{rg} 分别表示实验数据沿方位向和距离向的像元大小。

算法 3-1 基于 PCT 变换的视觉关注模型用于 SAR 图像舰船目标鉴别

算法 3-1 PCT based visual attention model for SAR ship discrimination

输入: 待检测影像 S 以及影像中的舰船目标候选框位置 $ShipProposal$;

输出: 优化后的检测结果 $RefinedShip$ 。

- 1: 计算待检测 SAR 图像中的候选目标数量 $N_p = size(ShipProposal, 1)$;
 - 2: 定义一个计数器 $count = 1$, 并初始化 $RefinedShip = zeros()$
 - 3: **for** $i \in [1, N_p]$ **do**
 - 4: 获取初始候选框位置 $[x_s^i, y_s^i, h^i, w^i] = ShipProposal(i, :)$;
 - 5: 获取自适应扩展候选框位置 $R_{ad} = [x_s^i - 2h^i + 1, y_s^i - 2w^i + 1, 5h^i, 5w^i]$;
 - 6: 通过 NLM 算法抑制相干斑噪声 $I_{NL} = NonLocalMean(R_{ad})$;
 - 7: 通过公式 (3-20) 对图像进行预处理并得到预处理后的图像块 I_{PP} ;
 - 8: 通过公式 (3-9) 至 (3-17) 获取自适应扩展区的视觉显著图 $I_{saliency}$;
 - 9: 通过公式 (3-18) 计算二值图像 I_{bi} ;
 - 10: 通过图像形态学处理得到优化后的目标并记录数量 N_m ;
 - 11: **for** $j \in [1, N_m]$ **do**
 - 12: 获得当前目标在自适应扩展区域中的位置: $temp = [x_{rb}^{ij}, y_{rb}^{ij}, h_r^{ij}, w_r^{ij}]$;
 - 13: **while** $(30 < h_r^{ij} \cdot PS_{az} < 600) \wedge (30 < w_r^{ij} \cdot PS_{rg} < 600)$ **do**
 - 14: **if** 目标不在陆地区域 **then**
 - 15: 计算目标在大场景影像中的位置;
 - 16: $RefinedShip(count, :) = temp + [x_s^i - 2h^i, y_s^i - 2w^i, 0, 0]$;
 - 17: 更新计数器 $count = count + 1$;
 - 18: **end if**
 - 19: **end while**
 - 20: **end for**
 - 21: **end for**
 - 22: **return** $RefinedShip$ 。
-

3.4.3 实验结果与分析

本实验训练和测试的数据集为 Sentinel-1 舰船检测数据集，实验的硬件环境和评价指标与上节一致。实验对比算法采用改进的多尺度卷积神经网络、级联卷积神经网络、以及结合级联网络与脉冲余弦变换的方法。

在算法性能评估时，本实验选取 25 个尺寸为 200×200 像素大小的图像块，共 324 个真值参与评价。表 3-5 列出了在 Sentinel-1 实测 SAR 图像上的舰船目标检测算法评估量化指标。从表中的统计结果可以看到：

- 结合级联网络与脉冲余弦变换的方法取得了最高的 F_1 值；
- 与多尺度卷积神经网络方法相比，级联网络能够在一定程度上提高目标检测性能。

表 3-5 结合级联网络与脉冲余弦变换方法的量化检测结果

Table 3-5 Quantitative detection results by combining cascade network and PCT for SAR ship detection

方法	真值数量	正确检测	错误检测	精度	查全率	F_1 值
改进多尺度网络	324	260	26	0.9090	0.8025	0.8524
级联卷积神经网络	324	283	23	0.9248	0.8735	0.8984
结合级联网络与脉冲余弦变换的方法	324	316	18	0.9461	0.9753	0.9605

图 3-17 展示了结合级联网络与脉冲余弦变换的方法在 Sentinel-1 影像上的视觉检测效果图，图中红色矩形框显示了舰船目标的检测结果。其中，图 3-17(a-d) 为四幅大小为 200×200 的原始图像切片。为了体现算法的优越性，本节选取的前两个测试图像中具有目标背景为高海况现象，第三个测试用例为近港/岸舰船目标密集排布区域的情况，测试用例四为陆地区域中存在与舰船目标散射情况极为相似目标的情况。图 3-17(e-h) 分别为改进多尺度卷积网络的检测结果，图 3-17(i-l) 分别为级联卷积神经网络的检测结果，图 3-17(m-p) 分别为结合级联网络与脉冲余弦变换方法的检测结果。通过观察直观的视觉效果图可以得出以下结论：

- 结合级联网络与脉冲余弦变换的方法能够使得最终的检测框更加精确，而且能够减少漏检和虚警，从而实现 SAR 图像中舰船目标的精细化提取；
- 与多尺度卷积网络相比，级联卷积网络模型能够在一定程度上提高检测效果，但是检测框的精确度仍然不足，而且依然存在陆地虚警的问题。

SAR 成像过程中由于采样频率有限而引起的方位向模糊的问题往往会导致 SAR 图像中出现一些鬼影目标。图 3-18 展示了 SAR 图像中的鬼影现象对舰船目标检测带来的

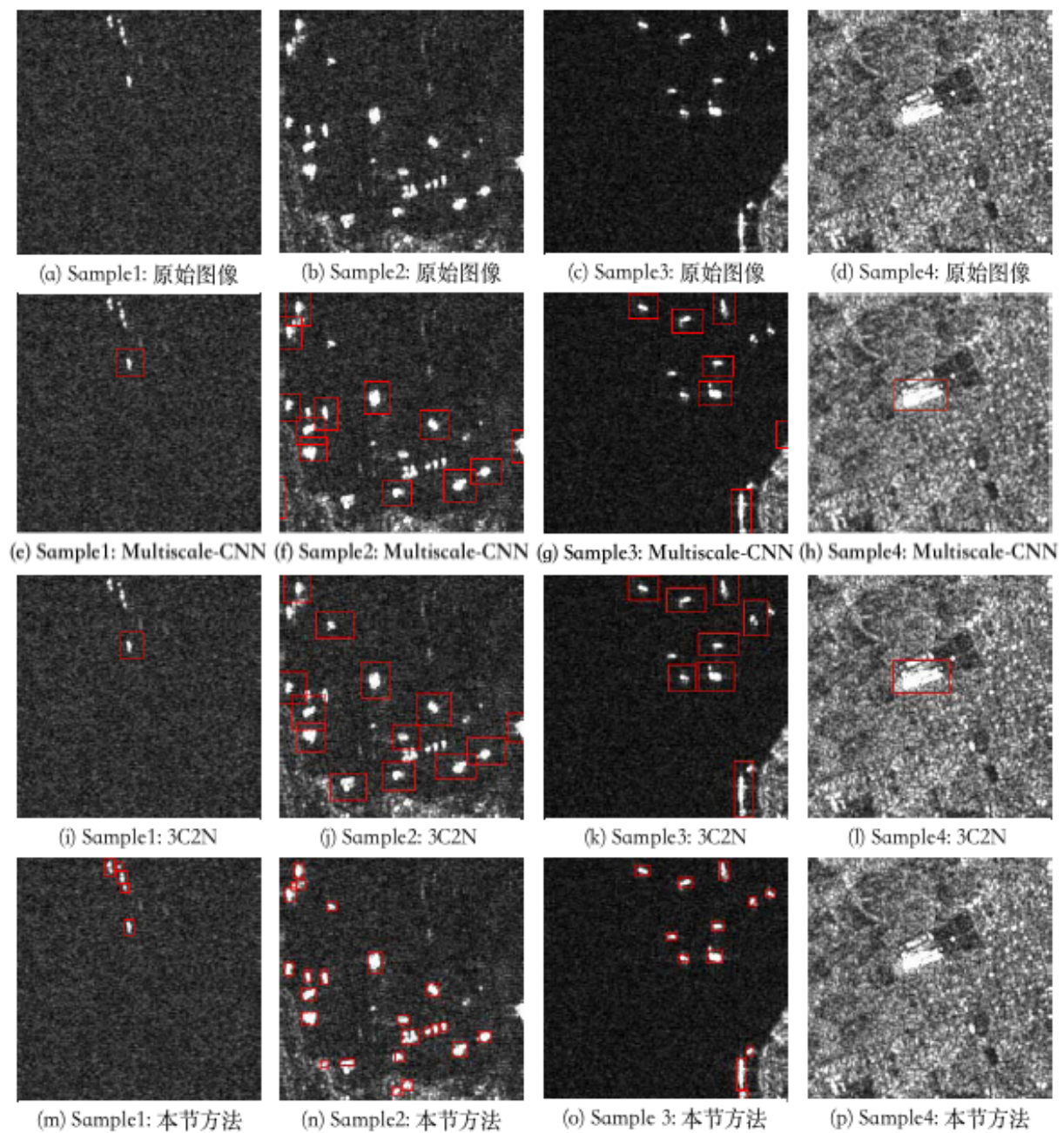


图 3-17 结合级联网络与脉冲余弦变换的方法在 Sentinel-1 实测数据上的检测效果图

Figure 3-17 Detection results on four image patches cropped from the Sentinel-1 images by combining cascade network and PCT

影响，图中黄色箭头连接了真实 SAR 舰船目标与其对应的鬼影目标。其中，图 3-18a 和图 3-18b 展示了级联卷积神经网络的检测结果；图 3-18c 和图 3-18d 展示了结合级联卷积神经网络和脉冲余弦变换的检测结果。从图中能够看到鬼影目标虽然与其对应的真实目标形状相似度高，但是其后向散射值却比较低，在真值标注的过程中，这些鬼影目标并没有被标记。因此，在检测结果中，这些虚假的鬼影目标并没有被网络检测到。这也说明了卷积神经网络方法是以数据为驱动的神经网络方法，对自身输入的真值数据及标注信息较为敏感，而在真值中没有被标记的鬼影也几乎没有被检测到。

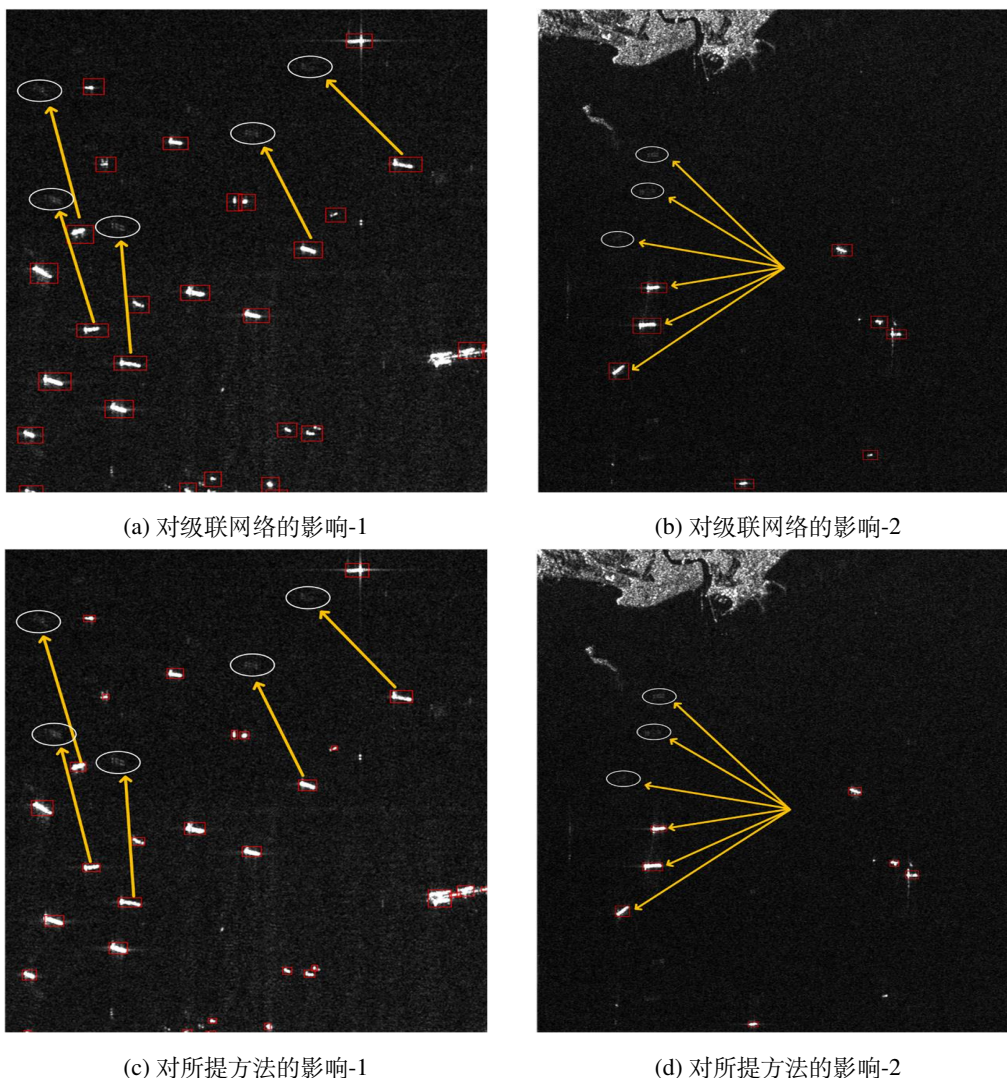


图 3-18 鬼影对舰船目标检测带来的影响

Figure 3-18 Goast analysis in ship detection

图 3-19所示为基于 PCT 的频域视觉模型处理的中间结果，图 3-19(a)-(i) 分别表示尺寸为 200×200 像素大小的原始 Sentinel-1 图像块、相干斑噪声抑制后的结果、预处理后的结果、图像平滑后的结果、二维 DCT 变换后的结果、频域处理后 IDCT 变换的结果、高斯滤波后的结果、形态学处理后的结果和经过舰船尺寸和位置约束后得到的结果。

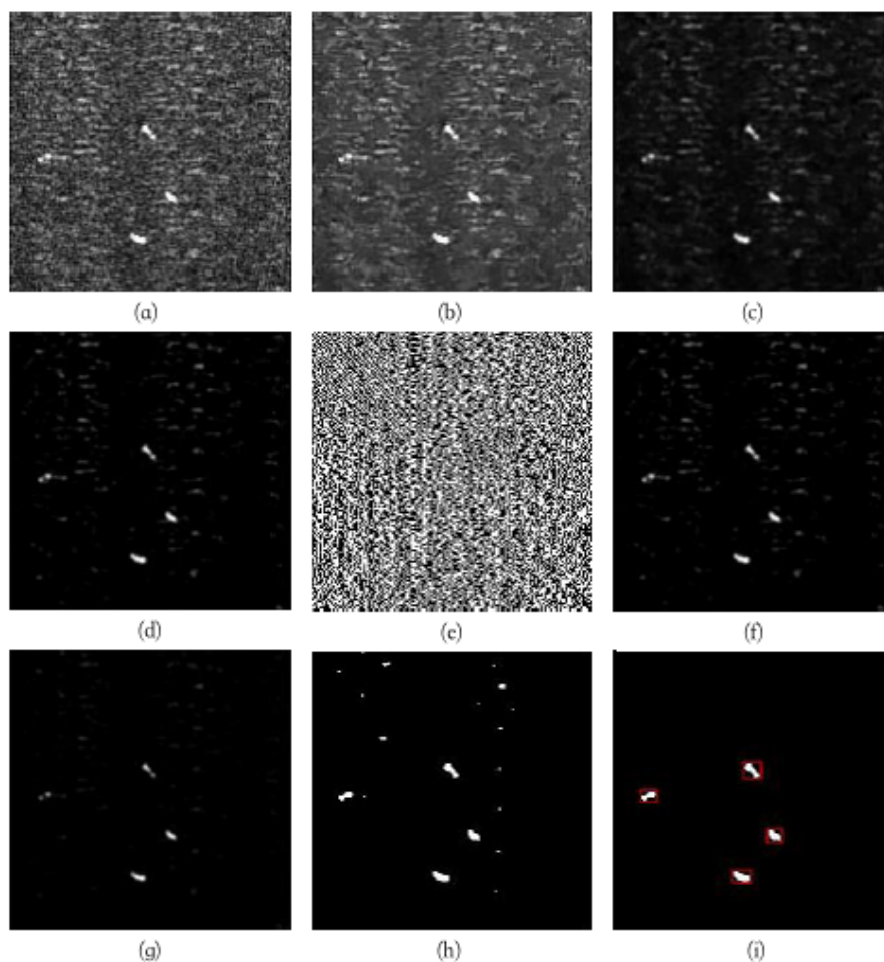


图 3-19 PCT 模型处理的中间结果。(a) 从原始 Sentinel-1 影像中截取的图像块；(b) 通过 NLM 算法降噪之后的图像块；(c) 图像预处理之后所得的结果；(d) 洪水图；(e) 频率图像块；(f) 逆变换回时域的图像块；(g) 高斯滤波后的图像块；(h) 形态学处理后的图像块；(i) 经过约束处理后的检测结果。

Figure 3-19 The intermediate results of PCT model. (a) one image patch tiled from the original Sentinel-1 images; (b) The image patch after image despeckling by using the NLM algorithm; (c) The image patch after preprocessing; (d) The flooding map; (e) The image patch in frequency domain; (f) The transformed image patch reversed to the space domain; (g) The image patch after Gaussian filtering; (h) The image patch after morphology; (i) The detection results by some constraints.

为了评价非局部均值滤波算法 NLM 算法用于 SAR 图像相干斑噪声抑制对最终舰船目标检测实验性能的影响,本节对比评估了相干斑噪声抑制对实验性能的影响。在本节所提出的结合级联卷积神经网络和基于脉冲余弦变换的视觉关注方法中未采用 NLM 算法做 SAR 图像相干斑噪声抑制时,所取得的检测结果如图 3-20(a-c) 所示。图 3-20(d-f) 展示了在本节方法中采用 NLM 算法抑制相干斑噪声后的检测结果。通过对比相应的实验效果图发现在缺乏相干斑噪声抑制过程时,检测结果中会存在较多虚警目标,而经过相干斑噪声抑制之后所得检测结果中检测准确性取得了一定的提升。

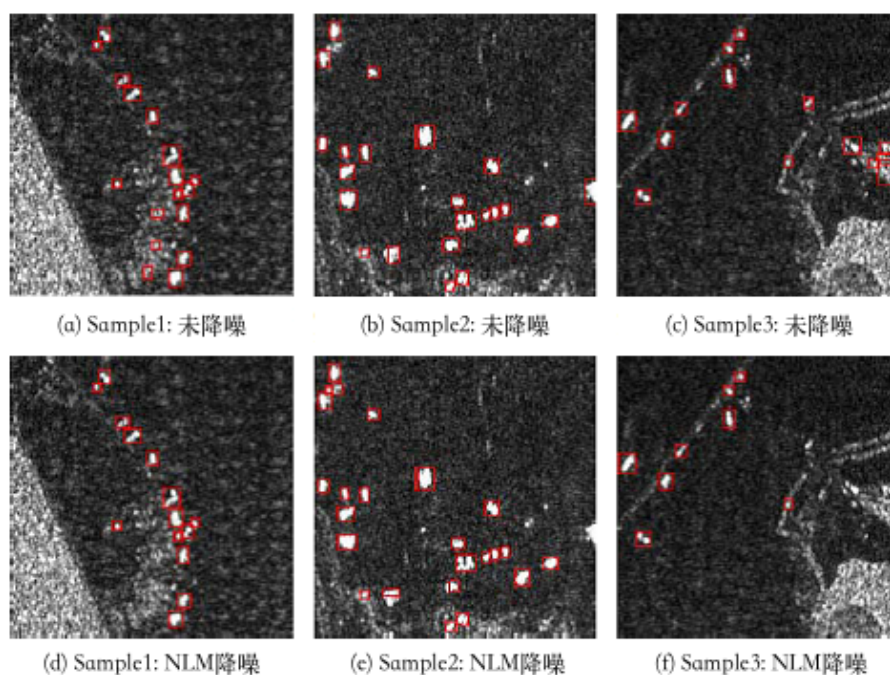


图 3-20 相干斑噪声抑制对检测结果的影响

Figure 3-20 Detection results with and without image despeckling

3.5 本章小结

本章实现面向 SAR 图像舰船检测的目标精细化解译。首先,本章突破传统多阶段检测算法的局限性,通过改进的多尺度卷积神经网络方法同时实现 SAR 图像中远海海域和进港/岸区域舰船目标的高精度检测。然后,针对多尺度卷积神经网络检测结果中存在的问题:1) 检测框精确度不足;2) 舰船分布密集区域存在较严重的漏检目标;3) 少量陆地虚警,同时考虑到卷积神经网络方法仅仅利用了图像的时域信息而忽略了图像的频域信息,本章通过结合级联卷积神经网络和基于脉冲余弦变换的频域视觉注意模

型，实现 SAR 图像中舰船目标的精细化提取，并分析了方位向模糊给检测结果带来的影响。在未来的研究中，可以研究通过一体化的深度网络框架实现时频信息联合学习。

第四章 多种极化方式下 SAR 图像的物理散射类型提取

4.1 引言

全极化 SAR 图像对地表覆盖的物理散射特性和几何特性均具有较强的可解释性。与之相反,单极化和双极化 SAR 图像由于缺乏全极化信息获取能力,对雷达物理散射特性的可解释性较为有限^[126-129]。研究者们已经针对全极化 SAR 图像,提出了多种极化目标分解方法,用于从全极化 SAR 图像中提取多种具有物理意义的散射类型,但是这些方法都难以适用于单极化和双极化 SAR 图像^[130]。

雷达物理散射特性是极化 SAR 图像解译的基础,并且在很多面向极化 SAR 图像的应用中都起着非常重要的作用,如城市规划^[131]、农业事件评估^[132, 133]、环境监测^[134]等。对于一个极化 SAR 系统来说,发射与接收电磁波的极化方式和电磁波波段是极其重要的两个参数,对极化信息等物理散射特性也具有非常深刻的影响。本章与下一章分别从多种极化方式和多种电磁波探测波段两个角度探讨雷达物理散射类型提取与特性分析,从而分别实现多种极化方式和多种探测波段下 SAR 图像地物散射特性的精细化解译。

本章从雷达极化方式的角度分析地物物理散射特性,重点从两个问题着手开展研究:如何从单极化和双极化 SAR 图像中提取蕴含在图像中的物理散射类型;与此同时,探索多种极化方式对 SAR 图像物理散射特性的影响。因此,本章的具体研究目标主要包括:1) 从单极化和双极化 SAR 图像中提取出具有物理意义的散射类型;2) 分析各种不同极化方式及其组合下的极化 SAR 图像对雷达物理散射特性的可解释性。

本章所研究的物理散射类型主要面向多次散射 (Multiple Scattering, MS)、体散射 (Volume Scattering, VS) 和表面散射 (Surface Scattering, SS)。多次散射主要发生在城区人造目标聚集区或者地表结构复杂的区域,如建筑物聚集的区;体散射则主要发生在森林地区、密集绿化带等区域;表面散射主要发生在平坦或粗糙的裸地、海上区域、河流等这样一些区域。图 4-1 简要形象地给出了本文所研究的物理散射类型的散射原理,包括多次散射、体散射和表面散射。

解决上述问题最直观的思路是首先从单极化和双极化 SAR 图像反演出其所对应的全极化 SAR 图像,然后通过极化目标分解的方式得到图像中具有物理意义的散射类型。但是,这种方法存在以下缺陷:1) 这种多阶段的方法容易存在误差层层传递而累积扩大的现象;2) 通过全极化反演的多阶段方法计算复杂繁琐。

因此,本章采用卷积神经网络方法通过端到端的学习方式,学习一种面向单/双极化 SAR 图像的具有物理可解释性的深度模型。所学习到的网络模型能够从单/双极化 SAR

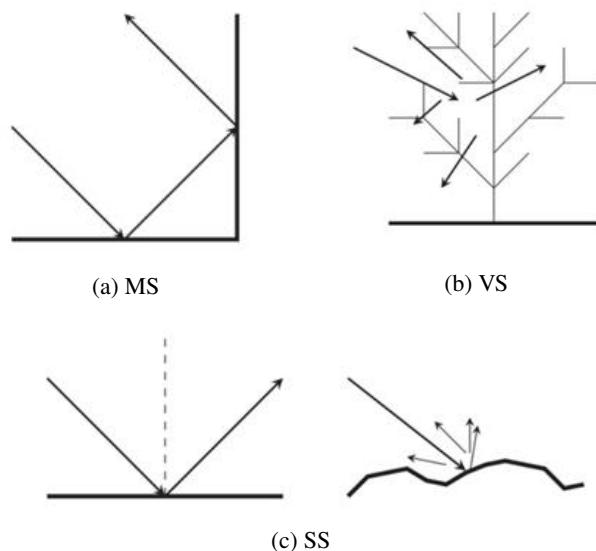


图 4-1 本文所研究的物理散射类型

Figure 4-1 Physical scattering signatures investigated in this study

图像中提取几种主要的物理散射类型，并对其性能进行全面分析评估。在设计网络模型的过程中，需要考虑到以下几点因素：1) 极化 SAR 图像原始的复数数据结构；2) 端到端一体化的模型结构；3) 模型具有较强的学习能力；4) 模型具有较强的泛化性能。综合考虑上述因素，本章采用基于复数域卷积神经网络的方法实现多种极化方式下 SAR 图像物理散射类型提取，从而从雷达极化方式的角度精细化解译 SAR 图像物理散射特性。训练所得的模型能够直接从原始的复散射矩阵中学习几种主要的物理散射类型。值得一提的是，在大多数情况下，研究者们为了简化问题往往只利用了幅度数据或者强度数据。但是，已有的相关研究表明，极化 SAR 图像中的相位信息在 SAR 图像解译中能够起到与幅度信息几乎相同的作用。

4.2 Cloude 极化目标分解

Cloude 极化目标分解方法是 S. R. Cloude 和 E. Pottier 于 1997 年提出的一种通过基于矩阵特征值分解的数学模型实现极化目标分解的方法^[88]，该方法基于二阶统计表征中的极化相干矩阵 T 通过一系列数学推导完成。由于极化相干矩阵 T 是一个 3×3 的共轭对称半正定矩阵，因此，极化相干矩阵 T 能够实现矩阵对角化，并且对角化之后的特征值均为正数、其对应的特征向量两两正交。相干矩阵特征值分解的数学表达式为

$$T = U \cdot \Sigma \cdot U^{-1} \quad (4-1)$$

其中,

$$\Sigma = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \quad (4-2)$$

是一个对角线元素由相干矩阵 T 的特征值组成的对角矩阵, 并且 $\lambda_1 \geq \lambda_2 \geq \lambda_3 > 0$ 。其对应的特征向量表示为

$$U = [u_1, u_2, u_3] = \begin{bmatrix} \cos \alpha_1 & \cos \alpha_2 & \cos \alpha_3 \\ \sin \alpha_1 \cos \beta_1 e^{j\delta_1} & \sin \alpha_2 \cos \beta_2 e^{j\delta_2} & \sin \alpha_3 \cos \beta_3 e^{j\delta_3} \\ \sin \alpha_1 \cos \beta_1 e^{j\gamma_1} & \sin \alpha_2 \cos \beta_2 e^{j\gamma_2} & \sin \alpha_3 \cos \beta_3 e^{j\gamma_3} \end{bmatrix} \quad (4-3)$$

其中, u_1 、 u_2 、 u_3 分别为特征值 λ_1 、 λ_2 、 λ_3 所对应的特征向量, 且三个特征向量两两正交, $(\alpha_1, \alpha_2, \alpha_3) \in [0, \frac{\pi}{2}]$ 为三个特征值分量所对应的目标散射角, $(\beta_1, \beta_2, \beta_3) \in [-\pi, \pi]$ 表示三个特征值分量所对应的目标方位角, $(\delta_1, \delta_2, \delta_3)$ 和 $(\gamma_1, \gamma_2, \gamma_3)$ 表示相应的相位角^[88]。

为了进一步分析地物要素的物理散射特性, Cloude 和 Pottier 在此基础上发展了三个新的极化参数, 用于后续判定物理散射类型。第一个参数为极化熵 (Entropy), 其数学定义为

$$H = - \sum_{i=1}^3 P_i \log_{\eta} P_i \quad (4-4)$$

其中,

$$P_i = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \lambda_3} \quad (4-5)$$

且 $i \in \{1, 2, 3\}$, 对数运算基由下式确定

$$\eta = \begin{cases} 3, & \text{单基站情况下} \\ 4, & \text{双基站情况下} \end{cases} \quad (4-6)$$

极化熵 H 表示散射目标由各向性散射至完全随机散射的随机性。

第二个参数极化各向异性 (Anisotropy) 用来衡量除主体散射之外的其余两个散射的强弱关系。该参数定义的表达式为

$$A = \frac{\lambda_2 - \lambda_3}{\lambda_2 + \lambda_3} \quad (4-7)$$

第三个参数为极化平均目标散射角 (α)。它描述了由表面散射至多次散射的平均随机性, 其定义如下

$$\alpha = \sum_{i=1}^3 P_i \alpha_i = P_1 \alpha_1 + P_2 \alpha_2 + P_3 \alpha_3 \quad (4-8)$$

在此基础上, Cloude 和 Pottier 给出了 $H - \alpha$ 极化分割平面, 该平面依据极化熵值和极化平均目标散射角的联合分布对地物目标的散射特性进行划分。 $H - \alpha$ 极化分割平面将极化熵 H 和极化散射角 α 分别作为横坐标和纵坐标, 在一个二维分布平面内将所有的散射类型共分为 9 个区间, 分布在每个区间的样本对应着不同的物理散射类型 (如图 4-2 中的 9 个数字所在的区间所示)。其中, 数字 1、2、3 所在的区间表示多次散射 (下文中分别用 Zone1、Zone2、Zone3 表示), 数字 4、5、6 所在的区间表示体散射 (下文中分别用 Zone4、Zone5、Zone6 表示), 数字 7、8、9 所在的区间表示表面散射 (下文中分别用 Zone7、Zone8、Zone9 表示)。Zone1 表示低极化熵多次散射区, 该区域中目标回波的极化熵较低, 能够代表各向同性的电解质和金属材质的二面角散射体等。Zone2 表示中等极化熵多次散射区, 该区域中回波的同极化分量间发生了 180 度方向偏移、多次散射和偶次散射, 通常城市区域、森林中的地面与树干存在的区域发生这种散射过程。Zone3 为高极化熵多次散射区, 该区域中散射回波在目标间多次反射, 能够代表粗壮的树木、建筑物等人造目标区。Zone4 表示低极化熵体散射区, 该区域中目标回波的同极化散射分量幅度差异较大, 一般能够解释具有较强各向异性的植被。Zone5 代表中等极化熵体散射区, 该区域中的目标也具有很强的各向异性, 多发生于茂密的植被覆盖区。Zone6 表示高极化熵体散射区, 该区域中的目标具有很强的各向异性, 能够代表具有各向异性的针状粒子的集合。Zone7 表示低极化熵表面散射区, 该区域中包含布拉格表面散射和镜面散射, 以及一些不会在同极化散射分量间引起 180 度方向偏移的散射过程。Zone8 代表中等极化熵表面散射区, 该区域目标具有较高的表面粗糙度。Zone9 表示高极化熵表面散射区, 该区域中一般存在具有表面散射机理且熵值特别大的散射体, 在现实情况中, 这种现象并不存在。表 4-1 更详尽地展示了 $H - \alpha$ 极化分割平面边界值和区域确定情况。

4.3 基于复数域卷积神经网络的物理散射类型提取

复数卷积神经网络 (Complex-valued Convolutional Neural Network, CV-CNN) 是复旦大学徐丰教授课题组于 2017 年提出的用于极化 SAR 图像地物分类问题的网络模型^[48]。CV-CNN 网络以全极化 SAR 图像的二阶相干矩阵 T 作为输入, 在网络运算过程中通过复数数据与网络参数之间的大量复数运算, 实现全极化 SAR 图像中的典型地物分类。在全极化的机载 E-SAR 实测数据上的实验结果显示出其在地物分类性能上的优势, 验证了复数数据结构在极化 SAR 图像解译中的优越性。

本章的研究目标是实现从多种极化方式 SAR 图像中提取具有物理意义的散射类型, 并探索不同极化方式下 SAR 图像的物理散射特性, 从而从雷达极化角度实现面向 SAR 图像地物散射特性的精细化解译。考虑到原始极化散射矩阵的复数数据结构、一体化的

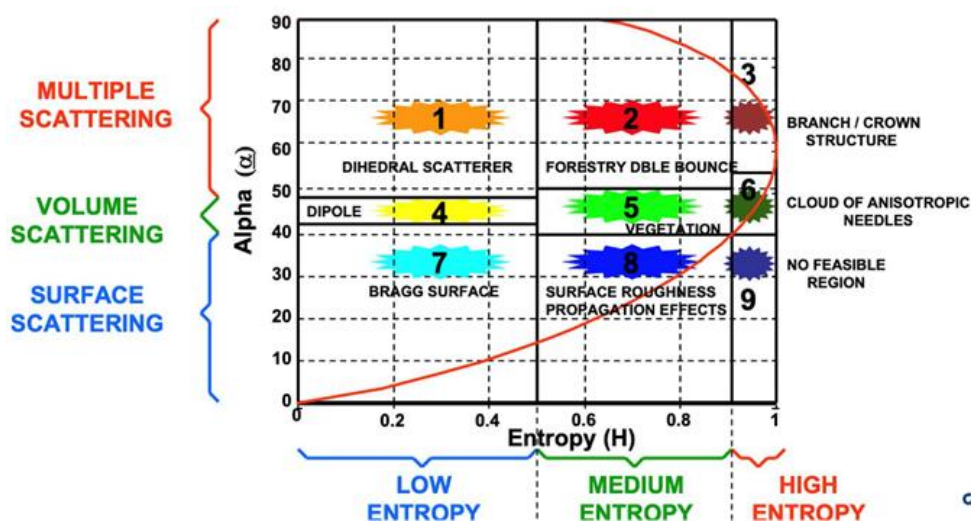
图 4-2 Cloude 的 $H - \alpha$ 极化分割平面Figure 4-2 Cloude's $H - \alpha$ polarimetric division plane

表 4-1 物理散射特性的真值分配方式

Table 4-1 Assignment of physical scattering types

极化物理散射类型 (数值标记)	参数范围
多次散射 (1)	$(0 \leq H \leq 0.5) \cap (47.5^\circ < \alpha \leq 90^\circ)$ $(0.5 < H \leq 0.9) \cap (50^\circ < \alpha \leq 90^\circ)$ $(0.9 < H \leq 1) \cap (55^\circ < \alpha \leq 90^\circ)$
体散射 (2)	$(0 \leq H \leq 0.5) \cap (42.5^\circ < \alpha \leq 47.5^\circ)$ $(0.5 < H \leq 0.9) \cap (40^\circ < \alpha \leq 50^\circ)$ $(0.9 < H \leq 1) \cap (40^\circ < \alpha \leq 55^\circ)$
表面散射 (3)	$(0 \leq H \leq 0.5) \cap (0^\circ \leq \alpha \leq 42.5^\circ)$ $(0.5 < H \leq 0.9) \cap (0^\circ \leq \alpha \leq 40^\circ)$ $(0.9 < H \leq 1) \cap (0^\circ \leq \alpha \leq 40^\circ)$

模型结构设计、模型需要具备较强的学习能力和泛化性能，本章采用基于复数域卷积神经网络的方法实现，网络通过优化复数域对比约束损失函数，从原始复数域 Sinclair 散射矩阵中学习参数，并输出具有物理意义的散射类型。

4.3.1 网络结构

基于复数卷积神经网络方法，本章从原始 Sinclair 散射矩阵 S 中直接学习蕴含在单极化和双极化 SAR 图像中的具有物理可解释性的模型，从而使得训练所得到的模型能够从单极化和双极化 SAR 图像中提取具有物理意义的散射类型，并在此基础上分析各种不同极化方式及其极化组合的 SAR 图像对于雷达极化物理散射特性的解译性能。本章所提出的复数域对比约束的卷积神经网络结构如图 4-3 所示。

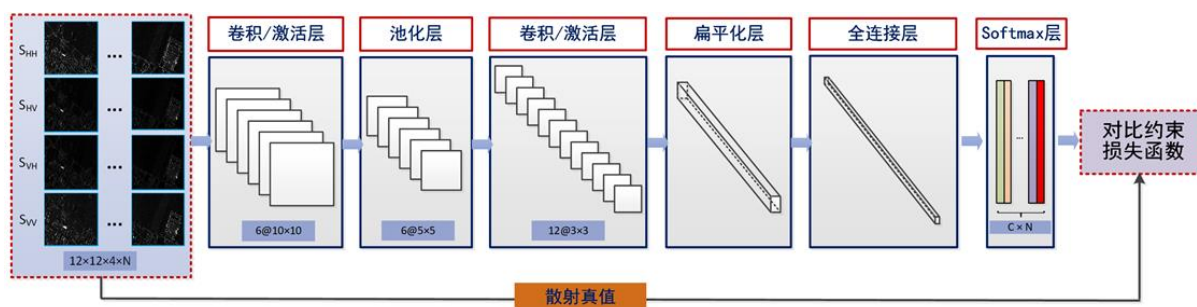


图 4-3 复数域对比约束卷积神经网络结构图

Figure 4-3 Network structure of the contrastive-regulated CNN in the complex domain

该网络接收极化 SAR 图像原始复数 Sinclair 散射矩阵作为输入数据，与网络中的复数参数进行大量复数域卷积、池化、激活等线性和非线性的运算，通过分批次梯度下降法对网络参数进行多次迭代更新。其中，复数域 softmax 运算的输出数据对应为各种雷达极化物理散射类型的类概率分布（实部和虚部均被认为是所预测的概率）。在复数域对比约束卷积神经网络运算中，卷积/激活层的运算规则在数学上表达为

$$\begin{aligned}
 O^{(l)} &= \Psi(O^{(l-1)} \otimes \Gamma^{(l)} + B^{(l)}) \\
 &= \{\Psi[\Re(O^{(l-1)}) \odot \Re(\Gamma^{(l)}) - \Im(O^{(l-1)}) \odot \Im(\Gamma^{(l)}) + \Re(B^{(l)})]\} + \\
 &\quad j\{\Psi[\Re(O^{(l-1)}) \odot \Im(\Gamma^{(l)}) + \Im(O^{(l-1)}) \odot \Re(\Gamma^{(l)}) + \Im(B^{(l)})]\}
 \end{aligned} \tag{4-9}$$

这里， Γ^l 和 B^l 分别表示第 l 个卷积层中的权重和偏置参数，

$$\Psi(x) = \frac{1}{1 + \exp(-x)} \tag{4-10}$$

表示 sigmoid 非线性激活运算， $\Re(\cdot)$ 和 $\Im(\cdot)$ 分别表示实部和虚部取值， $O^{(l-1)}$ 和 $O^{(l)}$ 分别表示第 $(l-1)$ 层和第 (l) 层的输出特征图。

复数域对比约束卷积神经网络中池化层操作分别在输入特征图的实部和虚部进行池化操作，其公式定义为

$$O_x^{(l)} = L_p(O^{(l-1)})_{x'} = L_p(\Re(O^{(l-1)}))_{x'} + jL_p(\Im(O^{(l-1)}))_{x'} \quad (4-11)$$

其中， L_p 表示在邻域 $x' = \Omega(x)$ 范围内的最大池化或平均池化，且 $x \in O^{(l)}$ ， $x' \in O^{(l-1)}$ 。

一般情况下，卷积神经网络在最后两层或者三层都是由全连接层构成。在全连接层中，每个神经元与前一层中每个神经元之间两两连接，具体的运算表达式为

$$O^{(l)} = O^{(l-1)} \cdot \Gamma^{(l)} + B^{(l)} \quad (4-12)$$

从式 (4-12) 可以看到，全连接层运算可以被看作是一种特殊的卷积运算。

网络的最后一层通常以一个 softmax 运算收尾，其输出为类概率分布，运算表达式为

$$\begin{aligned} O^{(l)} &= \text{softmax}(\Re(O^{(l)})) + j \cdot \text{softmax}(\Im(O^{(l)})) \\ &= \frac{\exp \Re(O_t^{(l)})}{\sum_{c=1}^C \exp \Re(O_c^{(l)})} + j \frac{\exp \Im(O_t^{(l)})}{\sum_{c=1}^C \exp \Im(O_c^{(l)})} \end{aligned} \quad (4-13)$$

这里， $\text{softmax}(\cdot)$ 为 softmax 运算且 $t \in \{1, 2, 3\}$ 。

4.3.2 网络优化的损失函数

在模型训练的过程中，本章期望网络的输出类概率能够无限制逼近其所对应的 one-hot 真值。本章中网络的优化损失度量主要基于两方面考虑：1) 本文期望通过常用的损失函数表达方式在复数域中重新对其定义，以达到基本的网络优化目的；2) 本文期望网络最后一层的输出类概率中，其类内样本的输出数据之间两两距离最小化并且在训练样本中类间样本两两之间的间距最大化。后者主要是考虑到极化 SAR 图像散射矩阵中同类散射类型散射系数差异大，而不同散射类型之间散射系数差异小的问题；以及散射真值获取过程中 $H - \alpha$ 极化分割平面的边界线附近样本难以有效区分的因素。

为了解决上述问题，本章在复数域卷积神经网络中的损失函数定义主要由复数域基础损失函数和对比约束项两部分构成，并通过二者加权求和得到最终的损失函数。因此，网络参数通过下式求最优解得到

$$\min_{\Theta} J(O^{(1)} | \bar{\mathcal{Y}}; \Theta) = \min_{\Theta} J_1(O^{(1)} | \bar{\mathcal{Y}}; \Theta) + \lambda J_2(O^{(1)} | \bar{\mathcal{Y}}; \Theta) \quad (4-14)$$

其中， J_1 和 J_2 分别表示所定义的复数域基础损失函数和对比约束损失项，加权系数表示为 λ 。在网络的训练过程中，最优参数通过最小化损失函数和反向传播机制，使得网络参数不断迭代更新，直至收敛。加权系数 λ 控制基础损失函数和对比约束正则项的相对重要性。

4.3.2.1 复数域基础损失函数

本文首先定义了复数域均方误差 (Complex-Valued Mean Squared Error, CV-MSE) 损失、复数域平均绝对误差 (Complex-Valued Mean Absolute Value, CV-MAE) 损失和复数域交叉熵损失 (Complex-Valued Cross Entropy, CV-CE), 并采用实验论证的方式从中则有选取在本文物理散射类型提取中最合适的一种度量方式作为最终的基础损失项。下面展开介绍基础损失项。假设对于第 i 个训练样本, 如果给定对应的 softmax 层输出的类概率与相应的 one-hot 雷达物理散射类型真值分别为 O_i 和 Y_i , 预测值与真值之间的误差表示为

$$E_i = O_i - Y_i \quad (4-15)$$

则 CV-MSE 基础损失项的数学定义为

$$J_1^{\text{MSE}} = \frac{1}{2N} \sum_{i=1}^N \sum_{c=1}^C [\Re^2(E_{ic}) + \Im^2(E_{ic})] \quad (4-16)$$

其中, $c \in \{1, 2, 3\}$ 表示物理散射类型标签索引。

与 CV-MSE 有所不同, CV-MAE 通过在同一训练批次内对训练样本 softmax 输出与雷达物理散射 one-hot 真值之间的误差总和最小化实现参数更新。CV-MAE 损失的数学定义为

$$J_1^{\text{MAE}} = \frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C [|\Re(E_{ic})| + |\Im(E_{ic})|] \quad (4-17)$$

CV-CE 基础损失项定义为

$$J_1^{\text{CE}} = \frac{1}{2N} \sum_{i=1}^N \sum_{c=1}^C \{[\Re(Y_{ic}) \cdot \log(\Re(O_{ic})) + (1 - \Re(Y_{ic})) \cdot \log(1 - \Re(O_{ic}))]^2 + [\Im(Y_{ic}) \cdot \log(\Im(O_{ic})) + (1 - \Im(Y_{ic})) \cdot \log(1 - \Im(O_{ic}))]^2\} \quad (4-18)$$

在上述定义的复数域基础损失函数基础上, 本章通过式 (4-19), 并根据多次实验结果择优选取性能最佳的损失函数作为最终的基础损失项。

$$J_1 = \text{OPT}(J_1^{\text{MSE}}, J_1^{\text{MAE}}, J_1^{\text{CE}}) \quad (4-19)$$

4.3.2.2 对比约束损失项

对比约束损失项设计主要考虑到两方面因素: 1) SAR 图像的非主观性等固有特性带来的散射系数类内差异大而类间差异小; 2) 在 $H-\alpha$ 极化分割平面中, 分布在散射类型划分边界上的样本极易混淆。本章通过在复数域中定义基础损失函数的同时加入对比

约束项使得训练样本通过层层运算后所输出的同类样本能够映射到距离更近的子空间，而来自不同极化物理散射类型的样本输出在距离更远的子空间。

复数域中的对比约束正则项具体定义形式为

$$J_2 = \begin{cases} [D(x_p, x_q) - \epsilon_1]^2, & \text{If } Y_p = Y_q \\ [\max(0, \epsilon_2 - D(x_p, x_q))]^2, & \text{Otherwise} \end{cases} \quad (4-20)$$

这里， x_p 和 x_q 分别表示同一训练批次内第 p 个和第 q 个样本的最后一层输出， Y_p 和 Y_q 分别表示相应的雷达极化物理散射 one-hot 真值， $D(\cdot, \cdot)$ 表示两个样本在最后一层输出数据子空间内的距离度量方式， ϵ_1 和 ϵ_2 分别用于放松类内间距和类间间距度量的约束条件。

当 x_p 和 x_q 两个样本来自同一散射类型 (即 $Y_p = Y_q$) 时，该正则项通过使得样本之间的距离最小化来实现凝聚类内样本的目的。这里， ϵ_1 是一个非常小的正实数。反之，当 x_p 和 x_q 两个样本来自不同散射类型 (即 $Y_p \neq Y_q$) 时，该正则项通过使得样本间的距离最大化来实现分离不同散射类型样本的目的。当不同散射类型在映射子空间的距离小于 ϵ_2 时，该正则项给损失函数加惩罚项 $[\epsilon_2 - D(x_p, x_q)]^2$ ，同时也通过最小化该约束项使得类间间距最大化；否则，不给损失函数增加惩罚项。本文通过额外引入两个变量 ϵ 和 δ 来保证 $\epsilon_2 > \epsilon_1$ 。这里，参数之间的关系表达为

$$\epsilon_1 = \epsilon - \delta \quad (4-21)$$

和

$$\epsilon_2 = \epsilon + \delta \quad (4-22)$$

其中， $\delta > 0$ 。

任意两个训练样本在网络输出子空间中的距离表示为

$$D(x_p, x_q) = \|\Re(Net(x_p)) - \Re(Net(x_q))\|_P + \|\Im(Net(x_p)) - \Im(Net(x_q))\|_P \quad (4-23)$$

其中， $\|\cdot\|_P$ 表示输入矩阵在第二维方向上的 P 范数运算。通常情况下， P 取值为 1、2 或者无穷大。特别地，本章在样本间距度量中采用经验值二范数，即 $P=2$ 。

复数域中的对比约束正则化项不仅能够使得网络学习到表征能力更强的特征表达，而且能够通过最小化损失函数使得网络凝聚类内样本的同时分离类间样本。网络中的复值参数能够从全极化数据中通过最小化损失函数的形式与反向传播机制不断迭代更新。网络训练所得的模型能够从单极化和双极化 SAR 图像中提取到具有物理意义的散射类型。

4.3.3 物理散射类型提取

为了从单极化和双极化图像中提取物理散射类型,本文首先从全极化 SAR 图像中学习到了一个具有物理可解释性的深度网络模型。该模型能够以单极化和双极化 SAR 图像作为输入,从原始的 Sinclair 散射矩阵中提取所蕴含的物理散射类型(包括多次散射、体散射、表面散射)。该方法同时利用了深度卷积神经网络极强的端到端学习能力和泛化性能,以及极化 SAR 图像的复数域数据结构信息,使得学习到的模型能够通过有效的复数参数类型对物理散射机理进行表征与建模。算法 4-1 为具体的散射类型提取算法流程,该算法详细描述了如何自动从给定的单极化和双极化 SAR 图像中提取具有物理意义的散射类型。在此基础上,探索多种不同极化方式对雷达极化物理散射特性的影响。

4.4 实验结果与分析

4.4.1 实验数据与网络训练真值

4.4.1.1 数据收集与预处理

本章的实验数据来自德国宇航中心的机载 L 波段高分辨率 F-SAR 数据。F-SAR 系统是德国最新研发的机载高分辨率 SAR 成像系统,它结合了多极化、多频段和干涉测量特性。F-SAR 系统主要针对该单位更早提出的 E-SAR 系统的缺陷在多方面做了改进。F-SAR 系统能够同时获取五个雷达电磁波波段(X、C、S、L 和 P 波段)、四种极化方式(HH、HV、VH 和 VV)的机载高分辨率 SAR 数据。

本章采用了德国 Kaufbeuren 地区和 Traunstein 地区共两景全极化 SAR 数据。前者用于模型训练与测试,后者用于模型泛化性能测试。其中, Kaufbeuren 地区数据的成像是 2016 年 1 月 21 日,关于该数据的一些重要参数信息如表 4-2 所列。在本章的研究工作中,训练复数域卷积神经网络所用的 F-SAR 数据在方位向和距离向的分辨率分别为 0.60 米和 1.92 米,图像像元大小沿两个方向分别为 0.36 米和 1.20 米。为了展示图像的直观视觉效果和基本的地物要素信息,本章对实验中所涉及的 F-SAR 数据在方位向做了 3 视处理,而距离向的视数保持不变。这样,经过多视处理后的实验数据在方位向和距离向的像元大小分别为 1.08 米和 1.20 米。

图 4-4 上半部分展示了训练/验证/测试集的生成过程,从左至右依次为:多视处理、数据归一化到零均值单位方差、训练/验证/测试数据采样。这里归一化在全极化数据各个极化通道中分别进行操作是为了便于卷积神经网络的训练^[48]。给定原始的 Sinclair 散

算法 4-1 从单/双极化 SAR 图像中提取具有物理意义的散射类型

算法 4-1 Extracting physical radar signatures from single- and dual-polarimetric SAR images.

输入: 原始极化 SAR 图像的散射矩阵 S ; 预训练的复数域物理可解释性模型 Net;

输出: 散射类型分布图 I_{out} ;

- 1: 计算图像通道数 N_{ch} 和图像尺寸 $[Row, Col]$;
 - 2: **if** $dim(S) = 2$ **then**
 - 3: 输入为单极化 SAR 图像;
 - 4: $[Row, Col] \leftarrow size(S)$;
 - 5: $N_{\text{ch}} \leftarrow 1$;
 - 6: **else**
 - 7: 输入为双极化或全极化 SAR 图像;
 - 8: $[Row, Col, N_{\text{ch}}] \leftarrow size(S)$;
 - 9: **end if**
 - 10: 分别沿方位向和距离向以固定的间隔从图像散射矩阵 S 中采样得到一系列尺寸为 12×12 像素大小的图像块;
 - 11: 根据公式 (4-24) 的原理对采样所得数据集进行归一化操作;
 - 12: 得到新的维度为 $12 \times 12 \times N_{\text{ch}} \times N_{\text{sn}}$ 数据集;
 - 13: 准备网络输入张量 D_{in} ;
 - 14: **if** $N_{\text{ch}} = 1$ **then**
 - 15: $D_{\text{in}} \leftarrow repmat(D, [1, 1, 4, 1])$;
 - 16: **else if** $N_{\text{ch}} = 2$ **then**
 - 17: $D_{\text{in}} \leftarrow repmat(D, [1, 1, 2, 1])$;
 - 18: **end if**
 - 19: 提取相应的物理散射类型 $D_{\text{out}} \leftarrow \text{Net}(D_{\text{in}})$;
 - 20: **for** $i = 1 \rightarrow size(D_{\text{out}}, 2)$ **do**
 - 21: 找到最接近 $(1 + j)$ 的一项并记录类别标签;
 - 22: **end for**
 - 23: 根据原图像尺寸重新组织物理散射类型;
 - 24: 通过红绿蓝三种颜色分别可视化多次散射 MS、体散射 VS 和表面散射 SS, 并输出可视化效果图 I_{out} 。
-

射矩阵 S_{uv} ，则数据的归一化操作在数学上表示为

$$S_{uv}^{\text{norm}} = \frac{S_{uv} - \mu(S_{uv})}{\sigma(S_{uv})} \quad (4-24)$$

其中, $(\cdot)^{\text{norm}}$ 表示归一化操作, $u, v \in \{H, V\}$ 表示雷达发射和接收电磁波的极化方式, $\mu(\cdot)$ 和 $\sigma(\cdot)$ 分别表示图像的均值和方差。其中, 图像均值的计算方式可以表示为

$$\mu(S_{uv}) = \frac{1}{\text{Row} \times \text{Col}} \sum_{h_1} \sum_{h_2} S_{uv}(h_1, h_2) \quad (4-25)$$

其中, Row 和 Col 分别表示图像在方位向和距离向的像素个数, h_1 和 h_2 分别表示方位向和距离向的像素索引。图像方差的计算表达式为

$$\sigma(S_{uv}) = \sqrt{\frac{\sum_{h_1} \sum_{h_2} [S_{uv}(h_1, h_2) - \mu(S_{uv})][S_{uv}(h_1, h_2) - \mu(S_{uv})]^*}{\text{Row} \times \text{Col}}} \quad (4-26)$$

这里, $(\cdot)^*$ 表示逐像素共轭运算。

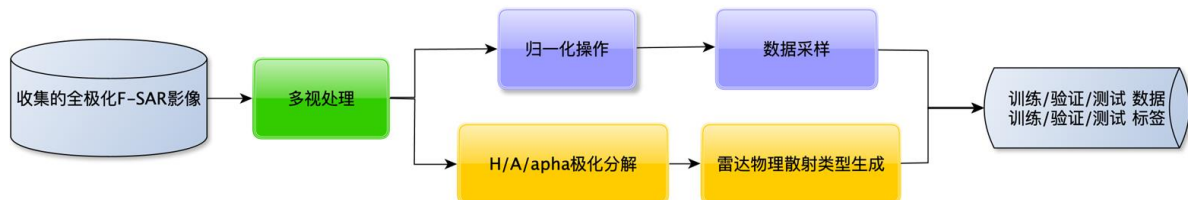


图 4-4 实验数据预处理与物理散射真值

Figure 4-4 Workflow of dataset preprocessing and ground-truth generation

研究表明图像在 12×12 像素大小的像素范围内能够表征地面目标特性^[48]。同时, 考虑到计算机的内存限制, 本章训练数据集的维度设置为 $12 \times 12 \times 4 \times N$ 。其中, N 为训练过程中同一批次的训练样本数量。

表 4-2 德国 Kaufbeuren 地区 F-SAR 影像基本参数信息

Table 4-2 Basic Information of F-SAR data acquired in Kaufbeuren, Germany

传感器	波段	成像时间 (年/月/日)	成像地点	图像大小 (az×rg/像素)	图像分辨率 (az×rg/米)	像元大小 (az×rg/米)
F-SAR	L	2016/01/17	德国 Kaufbeuren	90,112 × 2268	0.60 × 1.92	0.36 × 1.20

4.4.1.2 物理散射真值

本文基于 F-SAR 数据，通过 Cloude 极化分解得到极化熵 H 、极化平均散射角 α 和极化异向性值 A 。在此基础上，根据 $H - \alpha$ 极化分割平面划分情况得到的雷达极化物理散射类型图。图 4-5a 展示了实验数据的 $H - \alpha$ 参数分布情况；根据 $H - \alpha$ 平面划分后的雷达极化物理散射类型分布图如图 4-5b 所示。

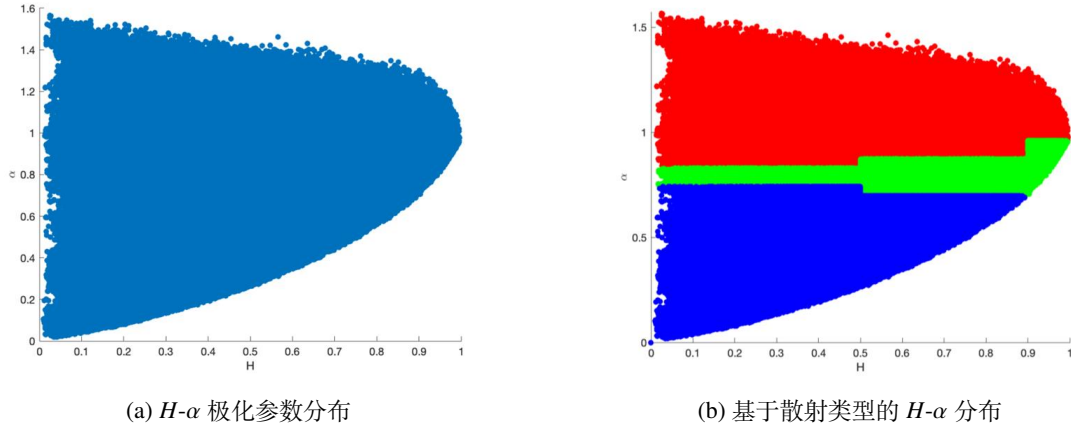


图 4-5 $H-\alpha$ 参数分布情况

Figure 4-5 $H-\alpha$ distributions

此外，对于训练/测试/验证数据集的划分，本章从实验数据中随机抽取约 60% 体量的样本用于模型训练同时确保训练样本数量为批次的整数倍，20% 的样本用做模型验证，剩余样本用于模型效果测试。因此，训练样本数量计算公式为

$$N_{\text{train}} = N \cdot \left\lfloor 60\% \cdot \frac{N_{\text{total}}}{N} \right\rfloor \quad (4-27)$$

测试样本数量为

$$N_{\text{test}} = \left\lfloor \frac{N_{\text{total}} - N_{\text{train}}}{2} \right\rfloor \quad (4-28)$$

验证集样本数量为

$$N_{\text{val}} = N_{\text{total}} - N_{\text{train}} - N_{\text{test}} \quad (4-29)$$

本章的研究目的是探索极化 SAR 图像中的多次散射、体散射和表面散射在不同极化方式下的可解释性能，并探索多种极化方式的 SAR 数据及其组合对这三种物理散射类型的解释性能。在真值获取方面，目前有多种极化目标分解方法都可以达到此目的。本章选择经典的 Cloude 分解，主要是考虑到以下两方面原因。第一，基于该极化分解方式而衍生的 $H - \alpha$ 极化分割平面直接为本文提供了获取地物物理散射真值的途径；第

二, 该方法计算量小且容易操作。基于 Cloude 分解和 $H - \alpha$ 极化分割平面而自动获取真值的方式在某种程度上使得整个网络学习过程可以实现无监督的学习方式, 从而有效地避免了昂贵的人工标注。 $H - \alpha$ 极化分割平面为雷达的物理散射真值获取提供了一种有效途径, 为卷积神经网络的训练提供了一种便捷有效的自动化获取手段。

图 4-4 下半部分展示了训练/验证/测试集中雷达极化物理散射真值的获取过程。通过 Cloude 分解获得的真值是逐像素的, 而对于给定的 $12 \times 12 \times 4$ 的局部范围来说, 本文选取出现频次最高的真值被简化为该局部范围内的真值。具体来说, 如果一个数值标记 $r \in \{1, 2, 3\}$ 在该给定的切片中出现的频次最高, 则该切片的真值标注为 r 。进一步地, 为了与卷积神经网络中最后一层 softmax 层的输出相吻合, 由上文的数值标记得到一个 one-hot 标记类型。所有样本的 one-hot 标记构成了一个 one-hot 标记集 $\bar{\mathcal{Y}} = \{Y_i; i = 1, 2, \dots\}$ 。这里, $Y_i \in \mathbb{C}^{1 \times C}$ 表示每个样本的 one-hot 标记, 且 $C = 3$ 是本章研究过程中的物理散射类型总数。值得一提的是, one-hot 标记向量中, 对第 y_i^{th} 个元素赋值为 $(1 + j)$, 其余元素赋值为 0。这里赋值为 $(1 + j)$ 的原因是在复数卷积神经网络中 softmax 层分别对实部和虚部进行操作, 因此该层输出的实部和虚部均被认为是类概率数值。

4.4.2 实验设置及性能评价

4.4.2.1 实验设置

在计算二阶统计量相干矩阵 T 的过程中, 本文设置了 3×3 、 5×5 和 7×7 三个不同大小的平均滑动窗。通过对比可视化实验结果发现, 采用 5×5 像素大小的滑动窗能够取得更加满意的效果。

在计算对比损失项的过程中, 本文设置参数 ϵ 和参数 δ 分别为 0.4 和 0.05。损失函数项的权重系数 λ 从 0 到 1 以 0.1 为步长寻找最优。

在网络训练过程中, 每个批次从训练集中随机抽取 100 个样本, 即 $N = 100$ 。初始学习率通过二分法多次训练寻优, 择优取学习率为 0.412。在训练过程中, 学习率以 0.01 的速率随训练批次下降, 动量参数设置为 0.99。网络参数更新策略采用随机梯度下降法不断迭代更新, 经过 40 次不断迭代更新直至损失函数收敛, 从而得到稳定的模型参数。

4.4.2.2 对比约束项对网络性能的影响

本章对对比约束项中通过加权系数均衡基本损失函数值和对比约束项对网络整体性能的影响。本文选取从 0 到 1 之间间隔为 0.1 值作为加权系数, 将网络验证集和测试集上的准确率为作为性能评价标准并从中选择最优参数。图 4-6 为所得到的结果示意图。

其中，红色虚线和蓝色虚线分别表示测试集和验证集上的准确率。本文发现当 $\lambda = 0.1$ 时测试集和验证集上的准确率能够达到最优。

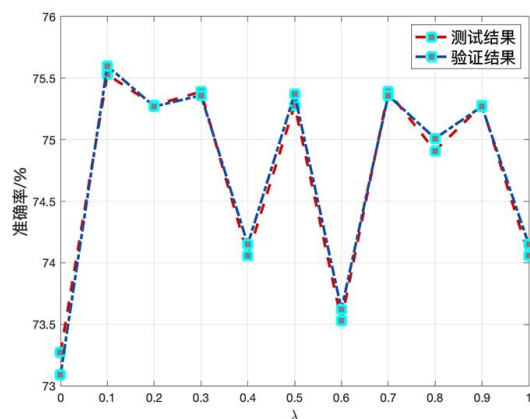


图 4-6 损失函数项加权系数的选取

Figure 4-6 Selection of the trade-off coefficient in overall loss measurement

4.4.3 多种极化方式下 SAR 图像的物理散射类型提取

本章通过模型训练所得到的复数域对比约束卷积神经网络模型能够实现从单极化、双极化和全极化 SAR 数据中自动提取蕴含在其中的雷达极化物理散射类型，从而能够探讨多种极化方式 SAR 图像对雷达极化物理散射类型的可解释性，最终从极化方式的角度实现面向 SAR 图像物理散射特性的精化解译。图 4-7展示了在各种不同极化方式数据及其组合下提取物理散射机制的准确率情况。其中，粉色折现表示整体评测的准确率，红、绿、蓝三种颜色的折线分别描述了多次散射、体散射和表面散射的提取准确率。参与评价的单极化图像包括 HH、HV、VH 和 VV 共四种极化方式的 SAR 图像；参与评价的双极化方式包括 HH 与 VV 结合 ([HH+VV])、HV 与 VH 结合 ([HV+VH])、HH 与 HV 结合 ([H*])、VH 与 VV 结合 ([V*])、HH 与 VH 结合 ([*H]) 与 HV 与 VV 结合 ([*V])，共六种不同的雷达极化组合方式。通过对比分析图 4-7所展示的结果，不难看出：

- HH 与 VV 两种共极化方式 SAR 图像结合能够显著提高雷达物理散射类型提取的整体准确率。
- 与单极化 SAR 图像相比，双极化模式对多次散射的解译性能均有不同程度的提升；
- 对体散射类型表现最好的极化方式为 HH 与 VV 结合的共双极化模式；

- 单极化 SAR 图像对表面散射类型具有较好的提取效果，但是极易将其他散射类型也误判为表面散射；
- 与单独的 HH 和单独的 VV 极化方式 SAR 图像提取结果相比，二者结合能够显著提高雷达物理散射类型提取的结果，而且与从全极化 SAR 数据中提取的结果相当。

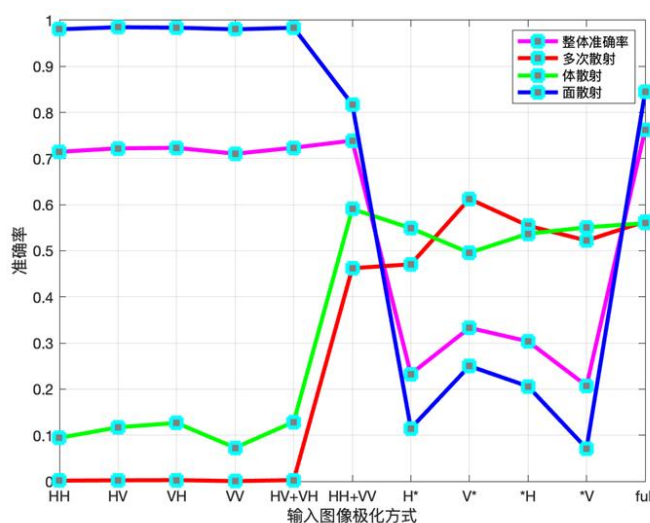


图 4-7 从各种极化方式 SAR 图像中提取物理散射类型的准确率比较

Figure 4-7 Accuracy comparison when extracting physical scattering signatures from different-polarimetric SAR images.

为了直观地展示所提取的物理散射类型分布情况，本章利用红、绿、蓝三种颜色分别表示多次散射、体散射和表面散射。图 4-8a 至图 4-8d 分别为 HH 极化、HV 极化、VH 极化和 VV 极化测试 SAR 数据通过对数变换可视化的图像。本文通过训练所得的复数域网络模型，从对应的 Sinclair 散射矩阵中直接提取物理散射类型，结果分别如图 4-8e 至图 4-8h 所示。从这些图中，本文发现单极化 SAR 图像对物理散射机制的解译能力非常有限，尤其是对多次散射现象的提取更是异常困难。

图 4-9a 可视化了通过 [HH+VV] 极化组合 SAR 图像所得到的物理散射类型。与单独的 HH 极化 SAR 图像提取结果和 VV 极化 SAR 图像提取结果相比，二者结合能够显著提高图像对物理散射机制的可解释能力。从图 4-9b 的 HV 结合 VH 极化 SAR 图像提取结果可以看到，HV 和 VH 结合的提取结果并没有非常明显的改善。图 4-9c 至图 4-9f 分别可视化了 [HH+HV]、[VH+VV]、[HH+VH] 和 [HV+VV] 极化结合方式的提取结果。与单极化 SAR 图像提取结果相比，这些结果都显示了其对雷达极化物理散射机制更强的可解释性。与此同时，[V*] 和 [*H] 极化结合方式对表面散射类型的解释能力较强，而

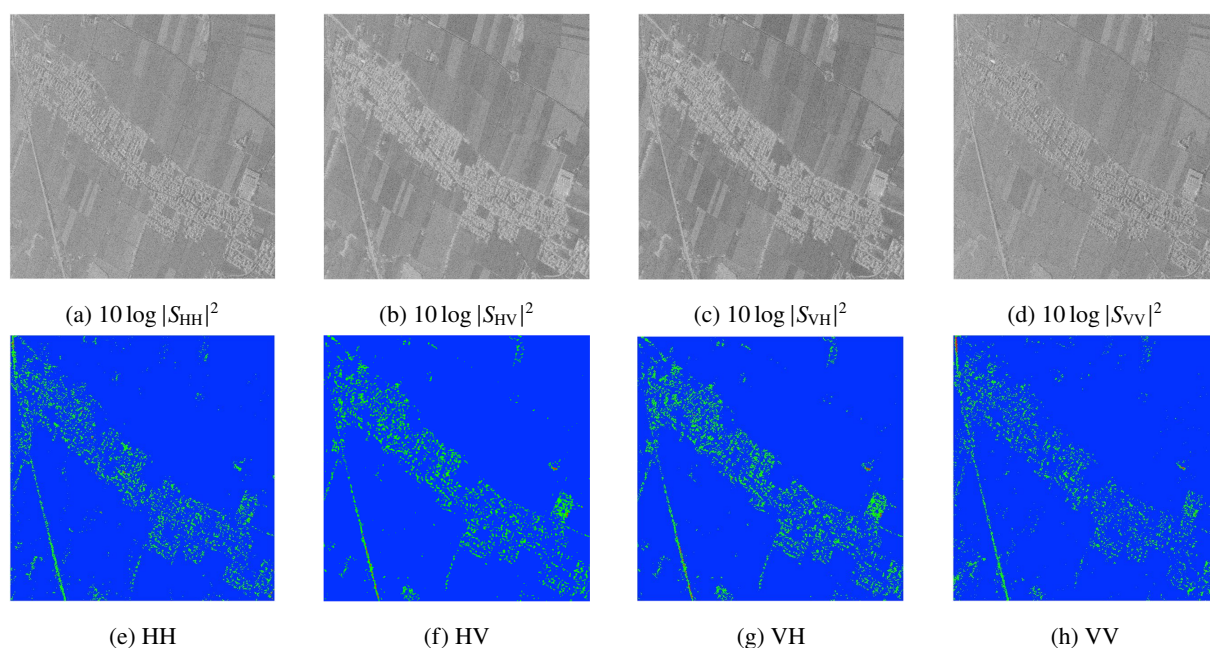


图 4-8 单极化 SAR 图像中提取物理散射类型的可视化结果

Figure 4-8 Visualization of physical scattering signatures from single-polarimetric SAR images

其他两种结合方式对表面散射的解释能力相对更弱一些。

4.4.4 验证模型的泛化性能

为了验证复数域对比约束卷积神经网络模型的泛化性能，本章的测试数据来源于 L 波段 F-SAR 数据。新的测试数据分辨率和像元大小分别为 0.60×1.29 米和 0.38×1.60 米 (azxrg)。在数据预处理过程中，本章首先对新的 L 波段数据沿方位向做了 4 视处理，并从中截取图像尺寸为 1000×1000 像素大小的数据块。为了可视化所选择的新测试数据，本文将其 Pauli 基可视化结果展示在图 4-10a 中。从这幅图中能够直观地看到成像地区的场景结构。图 4-10b 可视化了 Cloude 模型对全极化数据的物理散射机制分解结果。其中，红、绿、蓝三种颜色分别表示多次散射、体散射和表面散射。

图 4-11 展示了利用复数域对比约束卷积神经网络所训练的模型从不同极化方式 SAR 数据及其组合中提取物理散射类型的准确率，该图所得结论与图 4-7 所得结论保持一致。

图 4-12a 至图 4-12d 分别展示了 HH 极化、HV 极化、VH 极化和 VV 极化方式 SAR 数据模式下新的测试数据通过对数变换并可视化的结果。图 4-12e 至图 4-12h 为从相应的上一行数据中提取的雷达极化物理散射类型结果。

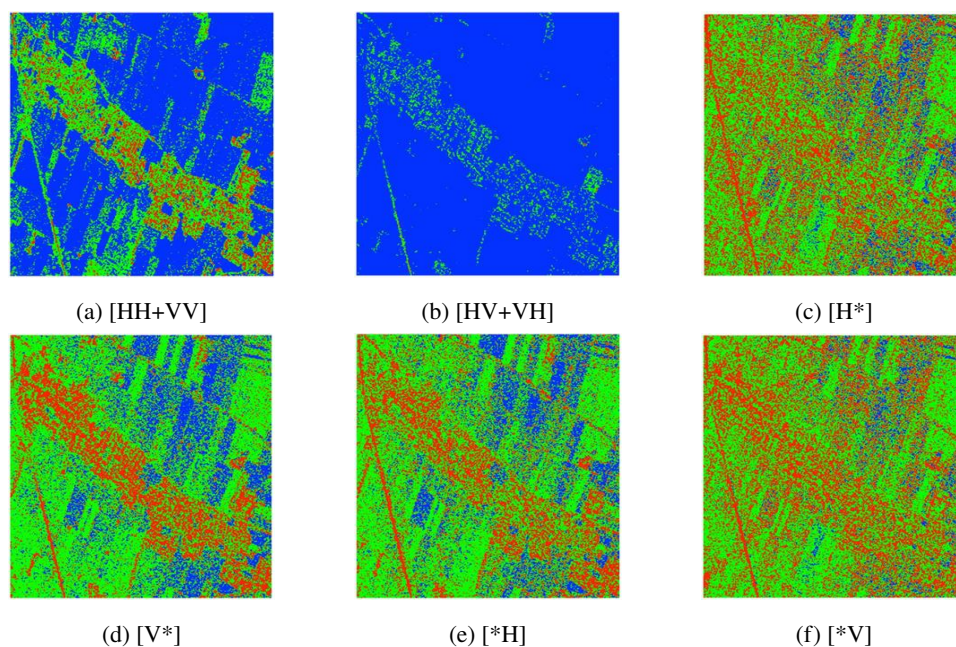


图 4-9 双极化 SAR 图像中提取物理散射类型的可视化结果

Figure 4-9 Visualization of physical scattering signatures from dual-polarimetric SAR images

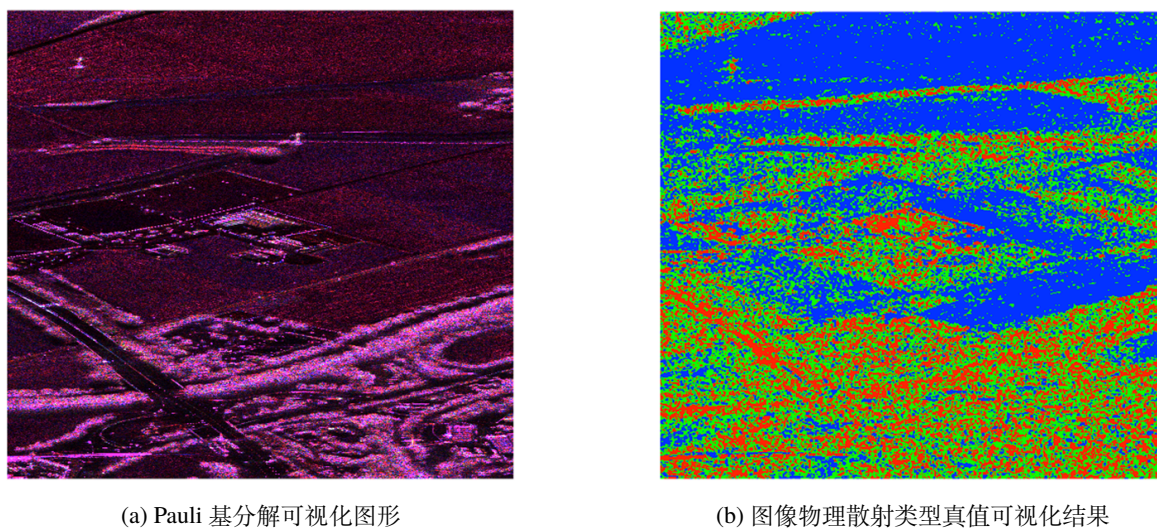


图 4-10 验证模型泛化性能所用的数据及其散射类型真值

Figure 4-10 Pauli visualization of the additional testing data and the ground-truth

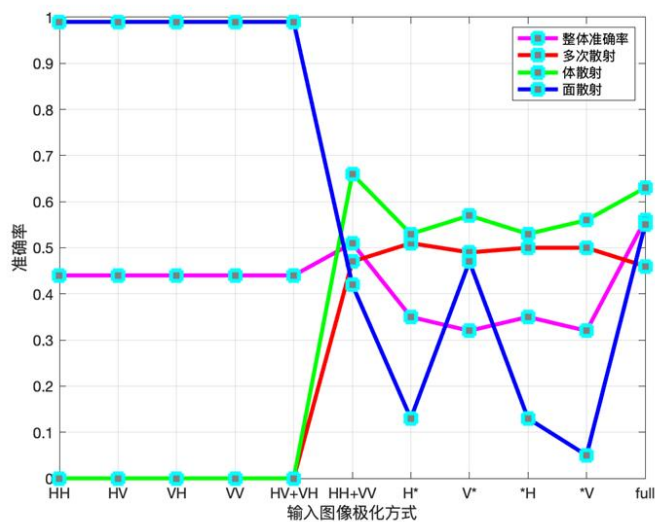


图 4-11 从多种极化方式新的 SAR 图像中提取物理散射类型的准确率

Figure 4-11 Accuracy comparison when extracting physical scattering signatures from newly multiple-polarimetric SAR images

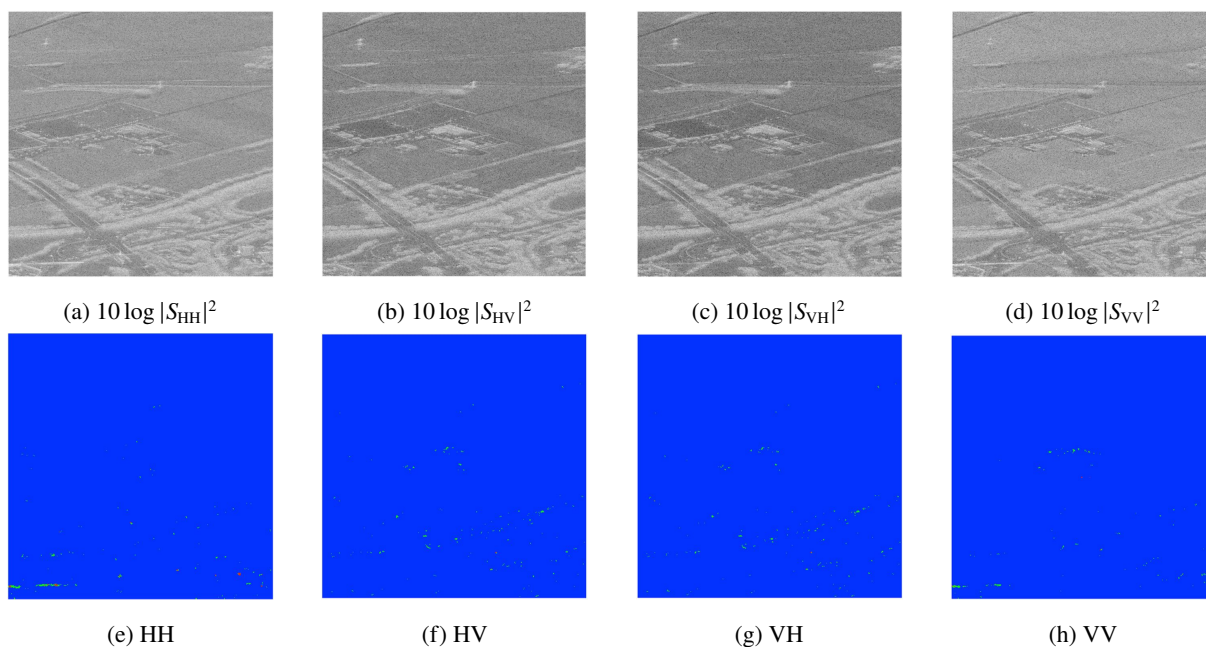


图 4-12 单极化 SAR 图像中提取物理散射类型的可视化结果

Figure 4-12 Visualization of physical scattering signatures from single-polarimetric SAR images

为了评价双极化组合模式下 SAR 数据对雷达极化物理散射类型提取的泛化性能,本章中图 4-13a所示为 HH 极化与 VV 极化方式组合 SAR 图像中雷达极化物理散射类型的提取结果,该结果再次验证了二者结合所组成的双极化 SAR 图像能够显著提高物理散射类型的解译能力。图 4-13b所示为 VH 和 HV 极化方式结合的双极化 SAR 数据中雷达物理散射类型的提取结果(与图4-9b的结论一致),该图所示结果也再次验证了交叉极化组合 SAR 数据对物理散射类型的提取非常有限。图 4-13c~ 图 4-13f的结果也表明了与单极化 SAR 图像相比,多种不同极化方式 SAR 图像组合能够在很大程度上提高模型对雷达物理散射特性的解译能力。

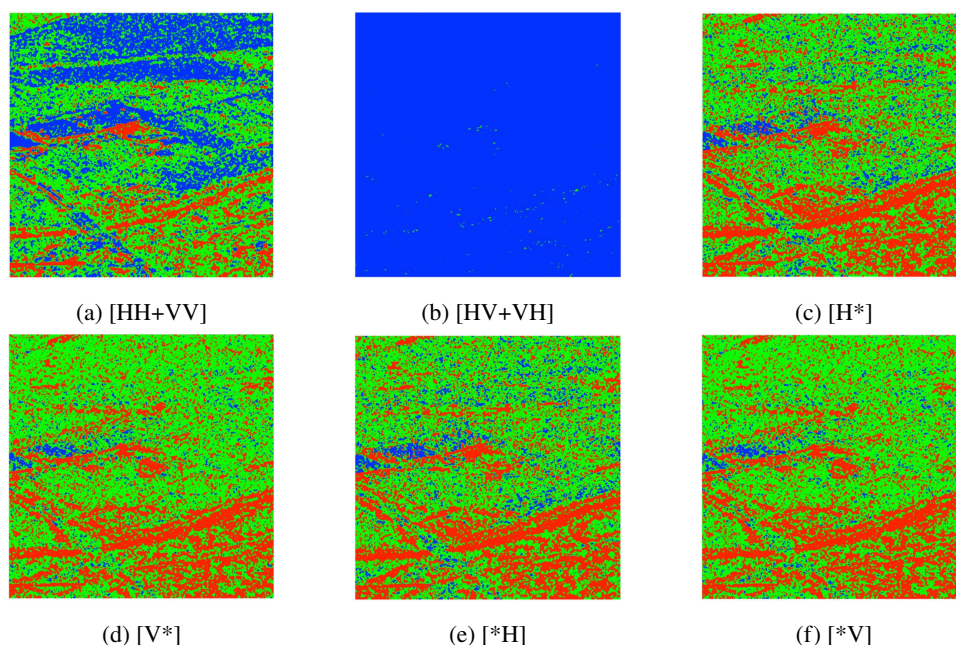


图 4-13 双极化 SAR 图像中提取物理散射类型的可视化结果

Figure 4-13 Visualization of physical scattering signatures from dual-polarimetric SAR images

4.5 本章小结

本章通过基于复数域对比约束卷积神经网络的方法实现了从单极化和双极化 SAR 图像中提取物理散射类型,并分析了多种不同极化方式 SAR 图像及其组合下的极化 SAR 图像对雷达物理散射机制的可解释性,从而从极化方式的角度实现精细化 SAR 图像散射特性解译。本章综合考虑了目前较为热门的深度卷积神经网络方法强大的学习与泛化性能和端到端一体化的训练机制,以及复数数据结构在极化 SAR 图像解译中的重要作用。复数域对比约束卷积神经网络模型能够直接从单/双极化方式 SAR 图像的

Sinclair 散射矩阵中提取到蕴含在极化 SAR 数据中的物理散射类型，并且比较各个极化方式及其组合的 SAR 数据对物理散射机制的可解释性能。在该方法中，网络训练的损失函数由复数域基础损失项和对比约束损失项通过加权求和得到，并监督网络的整个训练过程。对于基础损失项，本章定义了复数域均方误差、复数域平均绝对误差和复数域交叉熵损失，并根据在 F-SAR 数据上的实测结果择优选取基础损失项。对比约束项的设计目标是在增大训练样本的类间距离的同时缩小训练样本的类内距离。在多景机载 L 波段高分辨率 F-SAR 数据上验证了该方法的有效性。

第五章 多种探测波段下 SAR 图像的物理散射类型提取

5.1 引言

雷达极化物理散射类型是 SAR 图像解译中的基本要素之一，能够为多种极化 SAR 图像应用提供重要的基础。本章通过多种不同探测波段下 SAR 图像物理散射类型提取与分析，从雷达电磁波探测波段的角度实现面向 SAR 图像地物散射特性的精细化解译。一般而言，SAR 图像中的散射类型都是从全极化 SAR 图像中通过极化目标分解的方式获取。但是，由于全极化 SAR 图像的获取对成像技术具有较高的要求，代价也非常昂贵，在实际情况中，SAR 系统往往不具备全极化信息获取的能力。如欧空局的 C 波段 Sentinel-1 中 SAR 图像为双极化的成像模式，包含 VH 和 VV 两种极化方式；德国宇航中心的 X 波段 TerraSAR-X 图像为单极化成像模式，通常为 HH 极化方式。另一方面，这些缺乏全极化信息获取能力的 SAR 系统所成的图像通常具有较大的幅宽，如 IW 模式 Sentinel-1 影像的幅宽约为 250km。这种宽幅 SAR 图像在海洋、森林、农田等的连续监测上能够发挥重要作用。而具有全极化信息的 F-SAR 影像，其幅宽仅为 8.5km 左右。本章研究多种探测波段下 SAR 图像的物理散射类型提取。

为了从多种探测波段而不具备全极化信息的 SAR 图像中提取具有物理意义的散射类型，如 X 波段的 TerraSAR-X 影像和 C 波段的 Sentinel-1 影像等，一种有效的途径就是从 L 波段全极化 F-SAR 影像中通过复杂的非线性映射学习的方式获取到 X 波段和 C 波段影像中的物理散射类型。但是，要实现上述目的，本章需要考虑到以下几方面的难点问题。首先需要学习一个能够精确实现多种探测波段下 SAR 图像散射类型之间非线性映射关系的模型。其次，需要考虑到多种探测波段下 SAR 图像的成像几何特性和散射特性差异带来的影响。第三，需要考虑多种探测波段下 SAR 图像之间由成像时间和成像的地理位置带来的数据非对齐现象。

针对第一个难点，本章发现高斯过程回归方法能够提供一种有效解决该问题的途径。这是因为高斯过程回归方法在实现数据之间非线性映射的同时，不仅能够提供回归参数的预测均值，而且能够提供预测参数的置信范围。本章通过预测多种探测波段下 SAR 图像中的极化散射参数，进而推导出 SAR 图像中具有物理意义的散射类型。针对第二个难点，本章通过求期望和核函数样本关联的方式弱化多波段 SAR 图像中几何特性和散射特性不一致带来的影响。针对第三个问题，本章选取成像时间和成像地点均一致的多种探测波段 SAR 图像作为实验验证数据。

高斯过程假设每个训练样本服从高斯分布，任意多个样本服从多维高斯分布，通过

基于核函数的非线性映射从相应的散射数据中学习跨波段 SAR 图像中的极化参数，然后通过一定的推导方式得到 SAR 图像的物理散射类型，最终实现面向多种探测波段下 SAR 图像物理散射类型提取的精细化解译。

5.2 高斯过程基本原理及其在遥感图像中的应用

5.2.1 高斯过程基本原理

高斯过程假设单个像素的极化物理散射参数服从高斯分布，并且任意有限个像素的极化物理散射参数的联合分布服从多维高斯分布。定义 5.1 描述了高斯过程 (Gaussian Processes, GPs)^[135-137] 的基本定义。

定义 5.1 (高斯过程定义). 高斯过程是指一组随机变量的集合，且集合中任意单个随机变量服从高斯分布，任意有限多个随机变量均服从多维高斯分布。

高斯过程是一种可被用来对函数变量进行建模的数学模型，其关键思想是可以使用无限维多变量高斯分布来对函数进行建模。一个高斯过程由均值函数 $m(\cdot)$ 和协方差函数 $k(\cdot, \cdot)$ 确定。它天然地具有将有限的多维高斯分布拓展到无限维空间的能力。高斯过程由于所需调节的模型参数较少，因此，可以完美有效地结合先验知识与观测数据，模型预测结果也具有有良好的置信水平。高斯过程是一个基于函数分布的过程，它是建立在函数均值和方差基础上的。

假设目标变量可以通过一些未知的隐函数表征，这里假设为 $f(\mathbf{z})$ 。其中，输入训练数据集 $\{\mathbf{z}_n \in \mathbb{R}^D\}_{n=1}^N$ ，这里 N 和 D 分别为训练样本数量和特征维度。假设所需要预测的极化目标参数为 $\mathbf{y} = \{y_1, \dots, y_N\}$ ，这些数据集中的观测值可以看作是从一个联合高斯分布中得到的采样点，这样的一个数据集可以用作高斯过程建模。通常在假设联合高斯分布的均值处处为 0 的情况下，模型表达式为

$$\mathbf{y}_n = f(\mathbf{z}_n) + \varepsilon_n \quad (5-1)$$

其中，第 n 个样本的噪声参数服从零均值方差为 σ_n^2 的高斯分布，即 $\varepsilon_n \sim \mathcal{N}(0, \sigma_n^2)$ 。

一般地，假设隐函数 $f(\mathbf{z})$ 服从均值为零的高斯分布，且隐含噪声为 ε_n ，则高斯过程可以通过下式建模

$$f(\mathbf{z}) \sim \mathcal{GP}(m(\mathbf{z}), \mathbf{k}(\mathbf{z}, \mathbf{z}')) \quad (5-2)$$

其中，函数均值反映了所期望的隐函数在输入训练数据上的均值，它表示为 $m(\mathbf{z}) = \mathbb{E}[f(\mathbf{z})]$ 。通常情况下，为了防止运算复杂设置均值 $m(\mathbf{z}) = 0$ ，这一般通过在所有观测数据中减去均值来实现。协方差矩阵 $\mathbf{k}(\mathbf{z}, \mathbf{z}') = \mathbb{E}[(f(\mathbf{z}) - m(\mathbf{z}))(f(\mathbf{z}') - m(\mathbf{z}'))]$ 表示输入训练

样本 \mathbf{z} 和 \mathbf{z}' 之间的协方差，高斯协方差矩阵将训练与测试样本之间两两关联。核函数的选择与数据的分布形态有关^[138, 139]，其中一个最基本的约束是样本的距离与核函数大小成负相关的关系。一个广泛应用的核函数为

$$\mathbf{k}(\mathbf{z}, \mathbf{z}') = \sigma_f^2 \exp\left(-\frac{|\mathbf{z} - \mathbf{z}'|^2}{2\lambda^2}\right) \quad (5-3)$$

该核函数类型能够较好地对数据的连续性与稳定性进行建模，其中两个超参数 λ 和 σ_f^2 调节函数形状。一旦给定了均值函数与协方差函数，高斯过程即可被用于对有限的训练数据进行建模。

5.2.2 高斯过程在遥感图像中的应用

高斯过程回归方法作为一种有效的无参贝叶斯模型，通过在函数意义上的建模能够实现不同源数据集之间的参数映射，并且能够提供预测值的置信度。近年来，高斯过程在遥感图像物理参数预测中具有很多成功的应用案例^[140]。意大利 Yakoub Bazi 等人于 2008 年首次在高光谱图像分类中验证了高斯过程方法的有效性^[141]，随后在文献 [142] 中将高斯过程方法用于多光谱、多源、以及高光谱等遥感图像分类，并且验证了高斯过程用于多种遥感图像分类的性能优于支持向量机 (Supportive Vector Machine, SVM)^[143] 方法。之后，高斯过程模型被广泛应用于图像分类^[144]、语义标注^[145]、变化检测^[146, 147]、基于主动学习框架的遥感图像分类^[148, 149]。其中，西班牙瓦伦西亚大学 Gustau Camps-Valls 教授课题组做了大量基于高斯过程的遥感图像解译研究。Verrelst 等人就采用高斯过程方法通过将多个线性关系的光谱组合映射到高维空间实现非线性回归问题^[150]，从而实现农作物叶绿素含量、叶面积指数等参数的反演。在随后的几年时间里，这项技术被用在了基于 Sentinel-2 和-3 的生物与物理参数反演问题中^[151]。文献 [152] 的作者在所观测到的植被指数数值和可见光以及近红外图像中建立了一个非线性的模型用于未知参数的预测。Song 等在文献 [153] 中通过结合图模型框架^[154] 与高斯过程方法，巧妙地解决了高光谱图像分类中训练样本匮乏和高维度问题。除此之外，遥感图像解译领域中更多基于高斯过程的研究工作，如文献 [155-161]，也已经通过高斯过程回归模型成功地从遥感图像中预测到了所需的物理参数。此外，Katalin Blix 等人在文献 [162, 163] 中通过高斯过程方法在估计与海冰分析相关的参数，从而实现面向宽幅 SAR 图像的冰川图像监测。

5.3 基于高斯过程回归的物理散射类型提取

5.3.1 数据准备

本章研究所用到的参考数据来自德国宇航中心的高分辨率全极化 L 波段机载 F-SAR 数据, 图像尺寸沿方位向和距离向分别为 $90,112 \times 2,268$ 个像素。成像地点为德国 Kaufbeuren 地区 (地理位置大致在经度 10.62192 度, 纬度 47.88238 度), 成像时间为 2016 年 1 月 21 日。方位向和距离向的图像分辨率分别为 0.60 米和 1.92 米, 像元大小分别为 0.36 米和 1.20 米。本文对所研究的实验数据沿方位向做了 3 视, 而距离向的视数保持不变。因此, 多视处理后的图像像元大小在方位向和距离向分别为 1.08 米和 1.20 米。

对于全极化 SAR 图像来说, 本文能够借助多种极化目标分解方法, 如 Cloude 分解等, 从图像中提取具有物理意义的地物散射类型。相反地, 对于不具备全极化成像能力的 SAR 图像来说, 如 C 波段 Sentinel-1 图像和 X 波段 TerraSAR-X 图像, 由于缺乏全极化信息获取能力, 导致图像对地球观测的物理散射机制可解释性能严重受限。因此, 本章研究多种探测波段下 SAR 图像的物理散射类型迁移提取问题, 并且将物理散射类型聚焦在雷达多次散射、体散射和表面散射三种基本类型上。本章中所涉及的非全极化产品包括具备单视复数据后向散射获取能力的单极化 X 波段 TerraSAR-X 影像和双极化 C 波段 Sentinel-1 影像。本文基于参考数据 L 波段全极化 F-SAR 数据, 从源数据 TerraSAR-X 和 Sentinel-1 影像中 (本章中共涉及到两种不同的源数据类型) 提取具有物理意义的地物散射类型。

第一个源数据来自德国宇航中心的 HH 极化 Level 1B 级 TerraSAR-X 影像产品, 该数据是 2015 年 2 月 26 日获取的德国 Kaufbeuren 地区影像。该影像与 F-SAR 影像的成像时间差约为 11 个月。从谷歌地球的光学观测数据上观测得到该地区的影像没有发生较大变化; 另一方面, 由于源数据与参考数据的成像季节基本保持一致, 所以观测区域的地物后向散射类型基本保持不变。TerraSAR-X 源数据影像尺寸沿方位向和距离向分别为 $17,736 \times 6,940$ 个像素大小。影像沿方位向和距离向的分辨率分别为 1.12 米和 2.30 米, 像元大小分别为 0.17 米和 0.45 米。在图像预处理阶段, 本章对 TerraSAR-X 源数据沿方位向做了 5 视处理, 而距离向视数保持不变。因此, 在多视处理后, 图像对像元大小沿方位向和距离向分别变为 0.83 米和 0.45 米, 图像尺寸沿方位向和距离向变为 $3,547 \times 6,940$ 个像素。

第二个源数据来自欧洲空间局单视复数格式的干涉宽幅 (Interferometric Wide-swath, IW) Level-1 影像。需要指出的是 SLC IW 模式的 Sentinel-1 影像需要事先经过子带拼接得到宽幅整景 SAR 图像。该影像与 F-SAR 参考数据的成像时间差为 4 天, 这在很大程度上保证了源数据与参考数据之间的地物后向散射特性的稳定性。源数据 Sentinel-1 影

像共包括 VH 和 VV 两种极化方式,沿方位向和距离向的图像尺寸为 $13,118 \times 68,588$ 个像素。Sentinel-1 源数据沿方位向和距离向的像元大小分别为 13.9 米和 2.3 米。由于原始影像由三个子带数据构成,本章借助欧空局网站提供的 SNAP 工具包通过子带拼接的方式得到整个宽幅大场景图像。为了保证像元大小在方位向和距离向的基本一致性,本文对影像沿距离向做了 4 视,而方位向的视数保持不变。因此,多视处理后的图像像元大小分别为 13.9 米和 9.2 米,图像尺寸变为 $13,118 \times 17,147$ 个像素大小。更多的影像参数信息见表 5-1。

表 5-1 实验数据的基本成像参数

Table 5-1 Parameters of the dataset being used in this study

传感器	F-SAR	TerraSAR-X	Sentinel-1
电磁波波段	L	X	C
雷达极化方式	HH, HV, VH, VV	HH	VH, VV
图像分辨率 (方位向/米 \times 距离向/米)	0.60×1.92	1.12×2.30	5×22
图像像元大小 (方位向/米 \times 距离向/米)	0.36×1.20	0.17×0.45	13.9×2.3
多视处理 (方位向 \times 距离向)	3×1	5×1	1×4
图像尺寸 (方位向/像素 \times 距离向/像素)	$90,112 \times 2,268$	$17,736 \times 6,940$	$13,118 \times 68,588$
成像时间 (年/月/日)	2016/01/21	2015/02/26	2016/01/17

本章所讨论的具有物理意义的散射类型主要聚焦在多次散射、体散射和表面散射三种主要的散射类型上。图像中的物理散射类型基于预测的极化熵和极化平均散射角,通过 Cloude 分解理论中的 $H - \alpha$ 极化分割平面推导出对应的雷达极化物理散射类型。而全极化 F-SAR 数据的极化熵和极化平均散射角可以通过基于特征值分解的 Cloude 极化目标分解方法得到。

5.3.2 方法详介

为了较为准确地预测出相关的雷达极化参数,进而推导出所对应的物理散射类型,本章提出了一种基于核函数与平均期望的高斯过程回归方法。图 5-1 所示为该方法的具体实现流程图。图中训练所用的输入数据和训练后所得到的模型用黑色填充框表示,测试过程中的输入数据和所得到的物理散射类型用绿色填充框表示。图中展示了基于单极

化和双极化 SAR 图像的极化特征提取、基于全极化 F-SAR 影像的雷达极化物理散射类型获取、高斯过程回归模型及核函数选择、极化参数及其置信度预测等。特别地，紫色填充区域为极化 SAR 图像特征提取过程，橘色填充区域为 L 波段 F-SAR 数据的雷达极化物理散射类型的获取过程。该方法的核心思想是实现多种探测波段下 SAR 图像的极化参数非线性映射过程，从而根据一定的划分标准推导出所对应的雷达极化物理散射类型。在这个过程中，主要存在的挑战包括：来自于不同探测波段 SAR 图像的分辨率、电磁波波长、成像模式等参数之间存在较大差异，导致图像之间存在几何非对齐和散射非对齐的现象。本文所选取的源数据和参考数据在成像的地理位置和成像时间上尽可能地保持一致。该方法主要通过两个阶段的处理过程实现：第一阶段主要是极化 SAR 图像特征提取与物理散射类型的真值获取；第二阶段主要是完成模型训练与预测，以及所对应的物理散射类型推断。

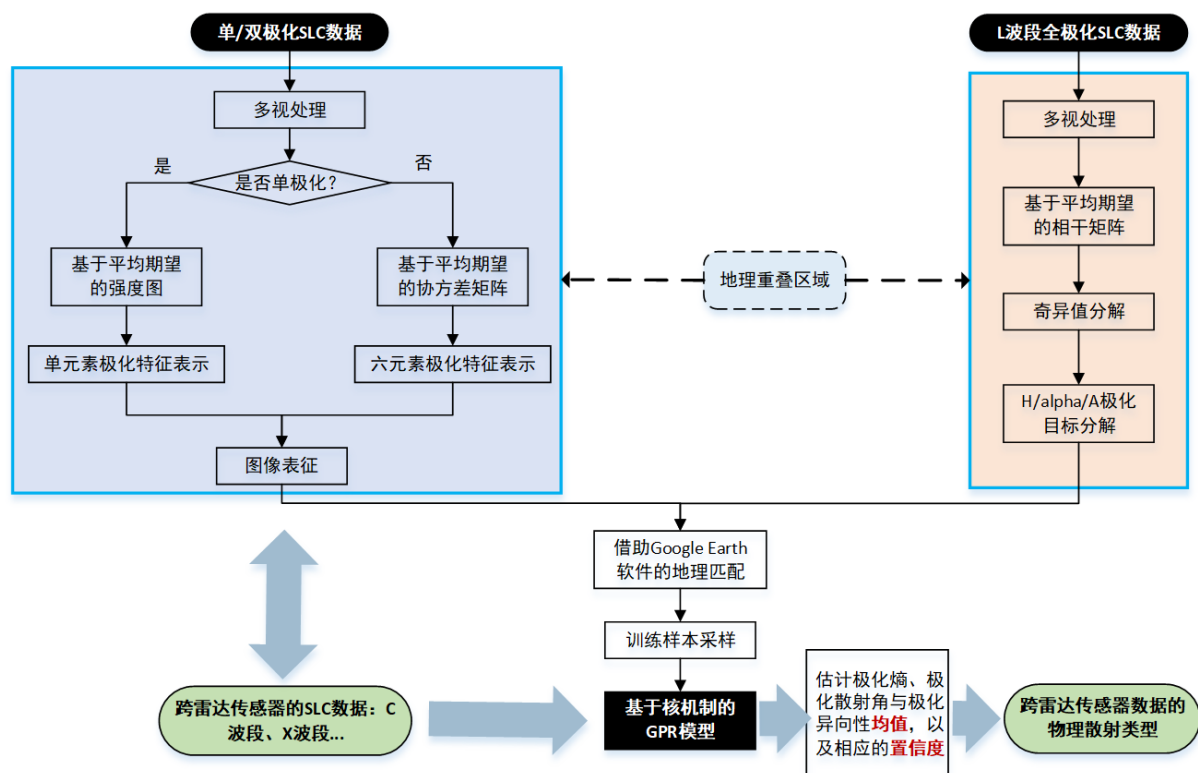


图 5-1 基于高斯过程回归的物理散射类型提取

Figure 5-1 Overall workflow of the proposed method

5.3.2.1 极化特征表示

在模型训练之前, 本文从 SLC 格式的单/双极化 SAR 图像中提取极化特征。对于 HH 单极化 TerraSAR-X 数据, 首先得到类协方差参数

$$C_{\text{tsx}} = E\{|S_{\text{HH}}|^2\} \quad (5-4)$$

这里, $|\cdot|$ 是指逐像素取绝对值运算, 下角标 tsx 表示 TerraSAR-X 影像。

基于该参数表征, 本章得到 HH 极化 TerraSAR-X 影像的极化特征, 其具体的数学计算公式可以表示为

$$\mathbf{z}_{\text{tsx}} = 10 \cdot \log C_{\text{tsx}} \quad (5-5)$$

其中, $\log(\cdot)$ 是取对数运算。

对于双极化源数据, 如 Sentinel-1 影像的 VH 和 VV 极化方式, 本文首先采用一个逐像素的 2×2 协方差矩阵表示图像中的每个像素, 具体的表达式为

$$C_{\text{sn}} = \begin{bmatrix} C_{11} & C_{12} \\ C_{12}^* & C_{22} \end{bmatrix} \quad (5-6)$$

其中, 下角标 sn 表示 Sentinel-1 影像, $(\cdot)^*$ 表示共轭运算, 且对角线变量可以表示为

$$C_{11} = E\{|S_{\text{VH}}|^2\} \quad (5-7)$$

和

$$C_{22} = E\{|S_{\text{VV}}|^2\} \quad (5-8)$$

非对角线元素表示为

$$C_{12} = E\{S_{\text{VH}}S_{\text{VV}}^*\} \quad (5-9)$$

然后, Sentinel-1 双极化源数据的极化特征可以表示为

$$\mathbf{z}_{\text{sn}} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \end{bmatrix} = \begin{bmatrix} 10 \cdot \log C_{11} \\ 10 \cdot \log C_{22} \\ C_{22}/C_{11} \\ |C_{12}|/\sqrt{C_{11}C_{22}} \\ C_{11} - C_{22} \\ (C_{11} - C_{22})/(C_{11} + C_{22}) \end{bmatrix} \quad (5-10)$$

5.3.2.2 问题描述与建模

为了从单极化或者双极化 SAR 图像中直接获取具有物理意义的地物散射类型，首先需要预测相应探测波段 SAR 图像中的极化熵和极化散射角。这些参数一般可以直接从全极化 SAR 图像中通过极化目标分解的方式得到。由于全极化 SAR 系统的成像过程对雷达成像能力的技术要求高、成像代价昂贵，全极化 SAR 图像，尤其是高分辨率全极化 SAR 图像在实际情况中并不常见。取而代之的是单极化和双极化 SAR 影像。如德国宇航中心的高分辨率 HH 极化星载 TerraSAR-X 影像和欧洲空间局的中低分辨率双极化 (VH 和 VV) Sentinel-1 影像。这些不同探测波段的数据，由于缺乏全极化信息获取能力，提取其中的雷达极化物理散射类型变得异常困难。本章研究从参考的 L 波段全极化 SAR 图像中通过参数非线性映射的方式，预测单极化和双极化 SAR 图像的极化参数信息，进而通过 $H - \alpha$ 极化分割平面得到其他探测波段 SAR 图像中的物理散射类型。

为了解决该问题，本文首先基于参考数据和源数据的地理位置和成像时间重叠区域，从源数据中通过高斯过程回归模型训练学习，预测得到 X 波段和 C 波段 SAR 图像中的极化参数信息（包括极化熵和极化散射角）。最后，本章基于所预测的极化参数，通过 $H - \alpha$ 极化分割平面得到最终的物理散射类型。在具体的模型构建过程中，极化熵 H 和极化平均散射角 α 均设置为高斯过程回归问题的目标参数 \mathbf{y} （见 5.2.1 部分）。

考虑到本章在研究过程中所涉及到的极化散射参数为连续数据，因而采用基于高斯过程回归 (Gaussian Process Regression, GPR) 的方法对所需的极化参数在源数据和参考数据之间建立非线性映射关系。高斯过程回归模型是一种专门用来解决基于高斯过程的回归问题，它是基于高斯过程理论发展起来的一种更加高级的机器学习方法，是核学习与贝叶斯推理相结合的产物，模型适应于处理高维度、小样本和非线性等复杂的回归问题，且具有较强的泛化能力^[164]。该模型假设任意单个变量服从高斯分布并且任意多个随机变量的联合分布服从多维高斯分布。高斯过程回归的优势可以总结为：易于实现、可以自适应获取超参数、能够灵活推断非参数、输出预测值的同时能够得到预测值的置信度。为了实现较为精确的极化参数预测，本章采用基于核函数的高斯过程回归方法，通过选择合适的核函数对训练数据和测试数据之间进行两两关联；另外，为了弱化几何和辐射非对齐带来的影响，本章在特征提取阶段采用求平均期望和核函数关联的操作缓和其对参数预测带来的影响。

5.3.2.3 高斯过程回归预测

假设给定含噪训练数据集 $\mathcal{D} = \{\mathbf{z}, \mathbf{y}\}$ ，本文期望能够通过后验概率 $p(f|\mathcal{D})$ 较为精确地预测到输入测试数据 \mathbf{z}_* 所对应的极化参数值 \mathbf{y}_* 。在高斯过程回归模型准备的过程中，所有数据对之间的关系可以通过以下几个协方差矩阵进行关联。它们是训练数据之

间的协方差矩阵

$$\mathbf{K}_{\mathbf{ff}} = \begin{bmatrix} k(\mathbf{z}_1, \mathbf{z}_1) & k(\mathbf{z}_1, \mathbf{z}_2) & \cdots & k(\mathbf{z}_1, \mathbf{z}_N) \\ k(\mathbf{z}_2, \mathbf{z}_1) & k(\mathbf{z}_2, \mathbf{z}_2) & \cdots & k(\mathbf{z}_2, \mathbf{z}_N) \\ \vdots & \vdots & \ddots & \vdots \\ k(\mathbf{z}_N, \mathbf{z}_1) & k(\mathbf{z}_N, \mathbf{z}_2) & \cdots & k(\mathbf{z}_N, \mathbf{z}_N) \end{bmatrix} \quad (5-11)$$

训练数据与测试数据之间的协方差矩阵

$$\mathbf{K}_{\mathbf{f}*}^T = \begin{bmatrix} k(\mathbf{z}_*, \mathbf{z}_1) & k(\mathbf{z}_*, \mathbf{z}_2) & \cdots & k(\mathbf{z}_*, \mathbf{z}_N) \end{bmatrix} \quad (5-12)$$

以及测试数据之间的协方差矩阵

$$\mathbf{K}_{**} = k(\mathbf{z}_*, \mathbf{z}_*) \quad (5-13)$$

通过核函数对比与实验论证, 本章选择自动相关决策 (Automatic Relevance Determination, ARD) 指数平方核函数来对数据进行两两关联。ARD 指数平方核函数的具体表达式为

$$k(\mathbf{z}_i, \mathbf{z}_j) = \sigma_f^2 \exp \left[-\frac{1}{2} \sum_{m=1}^D \frac{(z_{im} - z_{jm})^2}{\lambda_d} \right] \quad (5-14)$$

这里, λ_d 和 σ_f 是核函数中的两个超参数, λ_d 是数据集中两个样本点之间的距离对其相关性的影响程度, σ_f^2 描述了样本在纵坐标上的覆盖程度。从该式中也容易看出, 当 \mathbf{z}_i 与 \mathbf{z}_j 的距离越近时, 核函数的值越大, 而在预测新样本值的过程中, 距离当前样本点越远的样本其影响约小。

高斯过程回归方法是机器学习领域中的一种监督式函数逼近方法, 它利用样本点来捕捉数据的函数空间分布。训练与测试数据输出的联合分布可以通过以下高斯过程模型建模:

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{y}_* \end{bmatrix} \sim \mathcal{GP}(\mathbf{0}, \begin{bmatrix} \mathbf{K}_{\mathbf{ff}} + \sigma^2 \mathbf{I}_N & \mathbf{K}_{\mathbf{f}*} \\ \mathbf{K}_{\mathbf{f}*}^T & \mathbf{K}_{**} \end{bmatrix}) \quad (5-15)$$

这里, $(\cdot)^T$ 是矩阵转置操作。

定理 5.1 (高斯分布定理). 若两个集合变量 \mathbf{z}_1 和 \mathbf{z}_2 服从联合高斯分布, 则条件分布 $p(\mathbf{z}_1|\mathbf{z}_2)$ 服从高斯分布。

根据高斯过程理论可知, 如果两个集合变量 \mathbf{z}_1 和 \mathbf{z}_2 服从联合高斯分布, 则条件分布 $p(\mathbf{z}_1|\mathbf{z}_2)$ 服从高斯分布。由上述推导以及定理 5.1 可以得到模型预测输出的后验概率为

$$p(\mathbf{y}_*|\mathbf{z}_*, \mathcal{D}) \sim \mathcal{GP}(\mathbf{y}_*|\mu_{\text{GP}*}, \sigma_{\text{GP}*}^2) \quad (5-16)$$

其中, 预测均值可以表示为

$$\mu_{GP*} = \mathbf{K}_{f*}^T (\mathbf{K}_{ff} + \sigma^2 \mathbf{I}_N)^{-1} \mathbf{y} = \mathbf{K}_{f*}^T \boldsymbol{\beta} \quad (5-17)$$

预测值的方差可以表示为

$$\sigma_{GP*} = \sigma^2 + \mathbf{K}_{**} - \mathbf{K}_{f*}^T (\mathbf{K}_{ff} + \sigma^2 \mathbf{I}_N)^{-1} \mathbf{K}_{f*}, \quad (5-18)$$

从公式 (5-17) 的表达形式上看, 预测均值可以从两个不同的角度做出解释。一方面, 预测均值可以看作是 N 个核函数的线性加权和, 其中权重因子表示为 $\boldsymbol{\beta} = (\mathbf{K}_{ff} + \sigma^2 \mathbf{I}_N)^{-1} \mathbf{y}$ 。另一方面, 预测均值可以看作是训练数据输出 \mathbf{y} 的线性加权和。

5.3.2.4 模型参数优化

高斯过程回归方法中的参数主要集中在核函数中, 如尺度缩放因子 λ 、信号方差 σ_f^2 、噪声方差 σ_n^2 等。这些参数均为未知参数, 需要从数据中根据一定的算法原理进行估计。一般采用最大化对数似然的方法实现。假设超参数集合 $\boldsymbol{\theta} = \{\lambda, \sigma_f^2, \sigma_n^2\}$, 则相应的对数似然函数可以表示为

$$\log p(\mathbf{y}|\mathbf{z}, \boldsymbol{\theta}) = -\frac{1}{2} \mathbf{y}^T \mathbf{K}_y^{-1} \mathbf{y} - \frac{1}{2} \log |\mathbf{K}_y| - \frac{N}{2} \log 2\pi \quad (5-19)$$

其中, $\mathbf{K}_y = k(\mathbf{z}, \mathbf{z}') + \sigma_n^2 \mathbf{I}$ 是含噪输出 \mathbf{y} 的协方差函数, $-\frac{1}{2} \mathbf{y}^T \mathbf{K}_y^{-1} \mathbf{y}$ 项能够度量数据的拟合程度, $-\frac{1}{2} \log |\mathbf{K}_y|$ 能够度量惩罚项的计算复杂度, $-\frac{N}{2} \log 2\pi$ 是一个归一化项。

从而容易得到对超参数的梯度运算为

$$\begin{aligned} \frac{\partial}{\partial \theta_j} \log p(\mathbf{y}|\mathbf{z}, \boldsymbol{\theta}) &= \frac{1}{2} \mathbf{y}^T \mathbf{K}_y^{-1} \mathbf{y} - \frac{1}{2} \text{tr}(\mathbf{K}_y^{-1} \frac{\partial \mathbf{K}_y}{\partial \theta_j}) \\ &= \frac{1}{2} \text{tr}((\boldsymbol{\alpha} \boldsymbol{\alpha}^T - \mathbf{K}_y^{-1}) \frac{\partial \mathbf{K}_y}{\partial \theta_j}) \end{aligned} \quad (5-20)$$

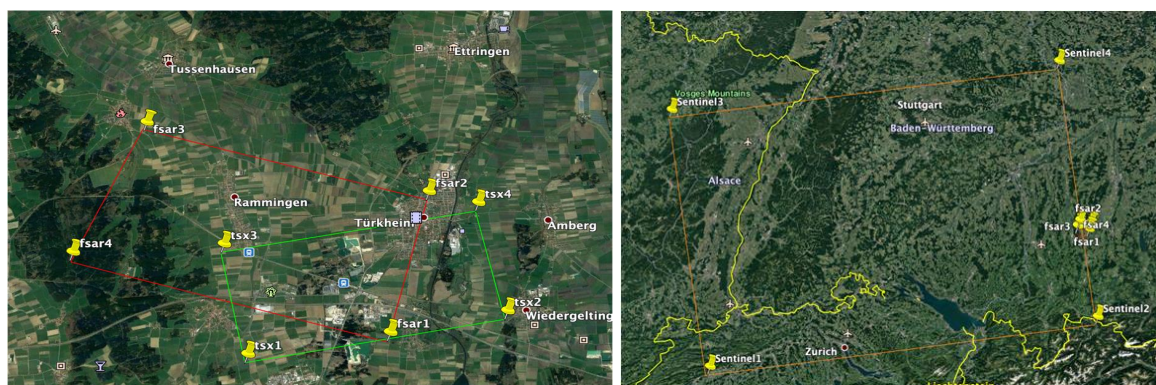
其中, $\boldsymbol{\alpha} = \mathbf{K}_y^{-1} \mathbf{y}$ 。

5.4 实验分析

5.4.1 实验设置

为了直观地展示所研究的区域, 本章利用谷歌地球光学数据通过地理定位的方式对实验区域进行展示 (如图 5-2 所示)。其中, 图 5-2a 显示了 F-SAR 数据与 TerraSAR-X 数据的覆盖范围, 图中红色框范围内为 F-SAR 影像的覆盖范围, 绿色框范围标注了 TerraSAR-X 影像的成像覆盖范围。F-SAR 数据与 Sentinel-1 影像的覆盖范围如图 5-2b 所

示，图中橙色和红色框所标注的范围分别是 Sentinel-1 影像和 F-SAR 影像的覆盖范围。从图中可以明显看到，F-SAR 影像与 TerraSAR-X 影像和 Sentinel-1 影像之间均有在地理位置上的重叠区域，这些重叠区域为本文提供了训练样本的采集来源。本章人为地在影像重叠区域标定了多次散射区域、体散射区域和表面散射区域，并针对每种散射区域分别采集 100 个样本点用于模型训练。相应的物理散射真值从全极化 F-SAR 图像中通过 Cloude 的极化目标分解和 $H - \alpha$ 极化分割平面得到。



(a) F-SAR 数据与 TerraSAR-X 数据的覆盖范围

(b) F-SAR 数据与 Sentinel-1 数据的覆盖范围

图 5-2 多种探测波段 SAR 数据覆盖范围

Figure 5-2 Training data overlaps of the two pairs of SAR images with multiple-bandwidth.

5.4.2 基于 X 波段 TerraSAR 数据的物理散射类型提取

对于参考数据 F-SAR 影像到源数据 TerraSAR-X 影像的物理散射类型分析，本章也是首先从多种探测波段的 SAR 图像中参数映射中预测源数据的极化物理散射参数，然后依据 Cloude 的 $H - \alpha$ 极化分割平面预测源数据中的物理散射类型。本章将在重叠区域中所采集的数据作为训练数据，具体是指从源数据中所获取的极化特征向量和从全极化 SAR 图像中获取的极化熵和极化平均散射角。然后，用这些采集的训练数据对训练基于核函数与平均期望的高斯过程回归模型。图 5-3 展示了 L 波段 F-SAR 影像和 X 波段 TerraSAR-X 影像的极化参数训练过程（包括极化熵和极化散射角），图中绿色边界灰色实心点表示原始训练数据，红色曲线为高斯过程回归拟合所得到的结果，蓝色虚线分别是预测参数的上边界和下边界。图 5-3a 和图 5-3b 分别是雷达极化熵和极化平均目标散射角的拟合结果。从图中能够看出，极化熵和极化散射角能够得到较好的拟合。

基于训练所得的高斯过程回归模型，本章从 TerraSAR-X 影像上非重叠区域中选取了一块地物类型较为丰富的图像尺寸为 600×800 像素大小的图像块进行模型效果测试。

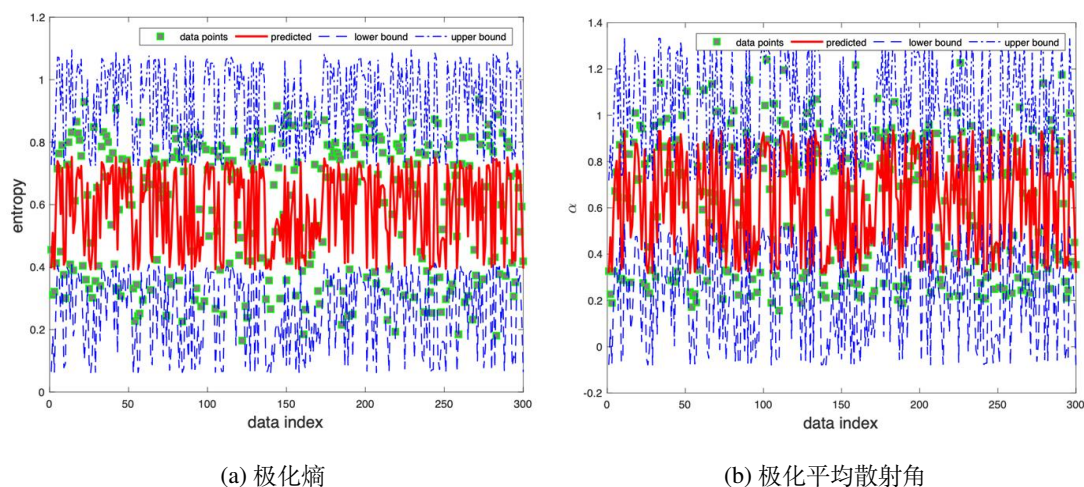


图 5-3 X 波段 TerraSAR 影像极化参数训练过程

Figure 5-3 Training process of polarimetric entropy and polarimetric alpha angle for X-band TerraSAR image.

图 5-4所示为对数变换后的 TerraSAR-X 图像块可视化，该图像的极化方式为 HH。从该可视化图中能够较为明显地看到图像中所覆盖地物的基本结构信息，这些复杂的地物结构信息恰好能够比较充分地验证模型的有效性。



图 5-4 通过对数变换的 TerraSAR-X 图像块可视化

Figure 5-4 TerraSAR-X test image after logarithmic transformation

图 5-5展示了在 X 波段 TerraSAR 数据上的所预测的极化熵参数的均值和置信水平。

其中, 图 5-5a 和图 5-5b 分别表示 TerraSAR-X 影像中极化熵的预测均值 \hat{H}^{tsx} 和预测值的置信水平 σ_H^{tsx} 。在这两幅图中以及下文的预测结果中, 高亮的颜色表示高的取值, 暗色表示相对较低的取值。

图 5-6 展示了在 X 波段 TerraSAR 数据上的极化平均散射角参数均值和置信水平。其中, 图 5-6a 和图 5-6b 分别表示 TerraSAR-X 影像中极化平均散射角的预测均值 $\hat{\alpha}^{\text{tsx}}$ 和预测值的置信水平 $\sigma_\alpha^{\text{tsx}}$ 。在图 5-6a 中高亮的颜色表示预测到的极化平均散射角值较大, 相反的情况预测到较低的极化平均散射角。图 5-6b 中的高亮值表示相应预测均值的不确定性较高, 反之表示预测的不确定性较低。

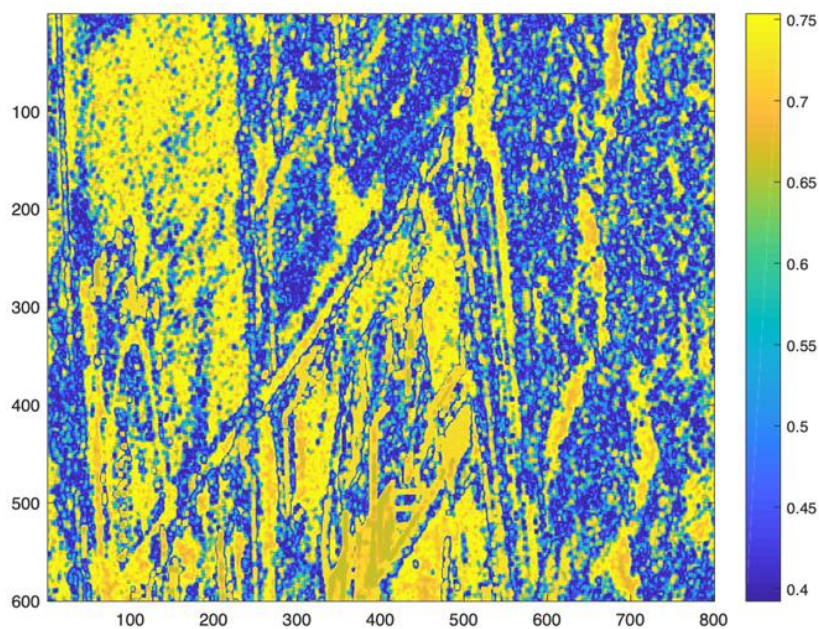
基于高斯过程回归模型所预测的极化熵均值 \hat{H}^{tsx} 和极化平均散射角 $\hat{\alpha}^{\text{tsx}}$, 本文依据 $H - \alpha$ 极化分割平面得到了 TerraSAR-X 影像的物理散射类型。图 5-7 所示为可视化的结果, 图中红、绿、蓝色分别表示图像中的多次散射、体散射和表面散射。这样, 本文就从单极化 X 波段的 TerraSAR-X 影像中得到了其所对应的物理散射类型信息。

5.4.3 基于 C 波段 Sentinel-1 数据的物理散射类型提取

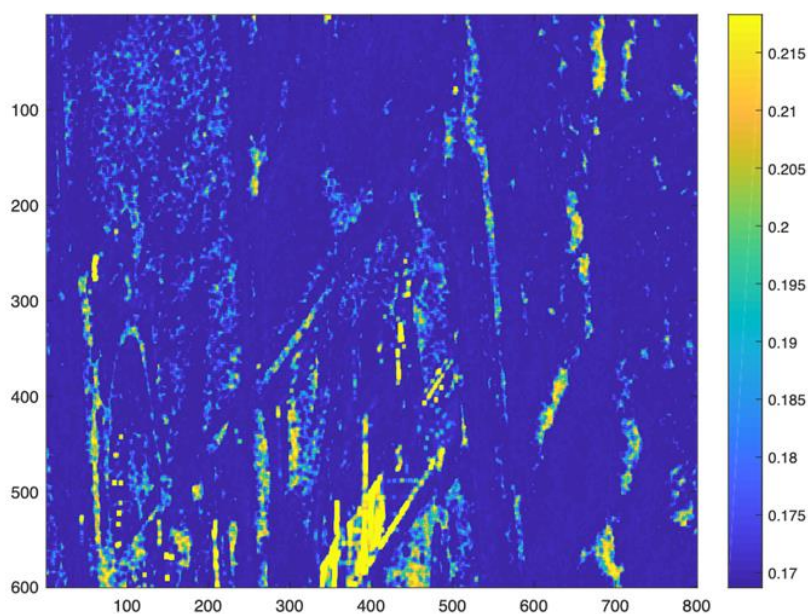
与 TerraSAR-X 影像上的实验类似, 本节在从 C 波段双极化的 Sentinel-1 影像中提取物理散射类型之前, 首先借助预训练的模型, 从 Sentinel-1 影像中预测极化物理散射参数。本文基于在 Sentinel-1 和 F-SAR 影像重叠区域中所采集的样本复数据和从全极化影像中获取的极化熵和极化平均散射角训练基于核机制和平均期望的高斯过程回归模型。图 5-8 展示了在 C 波段 Sentinel-1 影像上的极化参数训练过程, 图中绿色边界灰色实心点表示从重叠区域中所采集的原始训练样本, 红色曲线为高斯过程回归模型拟合所得到的结果, 蓝色虚线分别是所预测参数的上边界和下边界。图 5-8a 和图 5-8b 分别是极化熵和极化平均散射角的拟合结果。从图中能够看出, 极化熵和极化平均散射角能够得到较好的拟合。

为了评价该模型在 Sentinel-1 影像实测数据上的有效性, 本文从多视处理后所得的 Sentinel-1 影像上随机选取了一块 600×800 像素大小的图像块进行模型效果测试。图 5-9 所示为对数变换后的可视化 Sentinel-1 影像, 该影像共提供了 VH 和 VV 两种极化方式。其中, 图 5-9a 和图 5-9b 分别为 VH 和 VV 极化方式的 Sentinel-1 影像。所覆盖地物的基本结构信息能够较为完整地从小视化的 VH 和 VV 图像中观测到, 并能够观测到不同极化方式的 SAR 图像对不同地物的后向散射强度差异。

图 5-10 展示了从 C 波段 Sentinel-1 测试图像中所预测的极化熵参数的均值 \hat{H}^{sn} 和该参数所对应的置信水平 $\hat{\sigma}_H^{\text{sn}}$ 。其中, 图 5-10a 和图 5-10b 分别表示 Sentinel-1 影像中极化熵的预测均值 \hat{H}^{sn} 和预测值的置信水平 $\hat{\sigma}_H^{\text{sn}}$ 。在这两幅图中, 高亮的颜色均表示较高的取值, 暗色均表示相对较低的取值。在图 5-10a 中, 高亮的颜色表示预测的极化熵值



(a) 极化熵预测均值



(b) 极化熵预测值的置信水平

图 5-5 X 波段 TerraSAR 数据上预测的极化熵均值和相应的置信水平

Figure 5-5 Predicted mean value of polarimetric entropy and the corresponding certainty level from a TerraSAR-X image

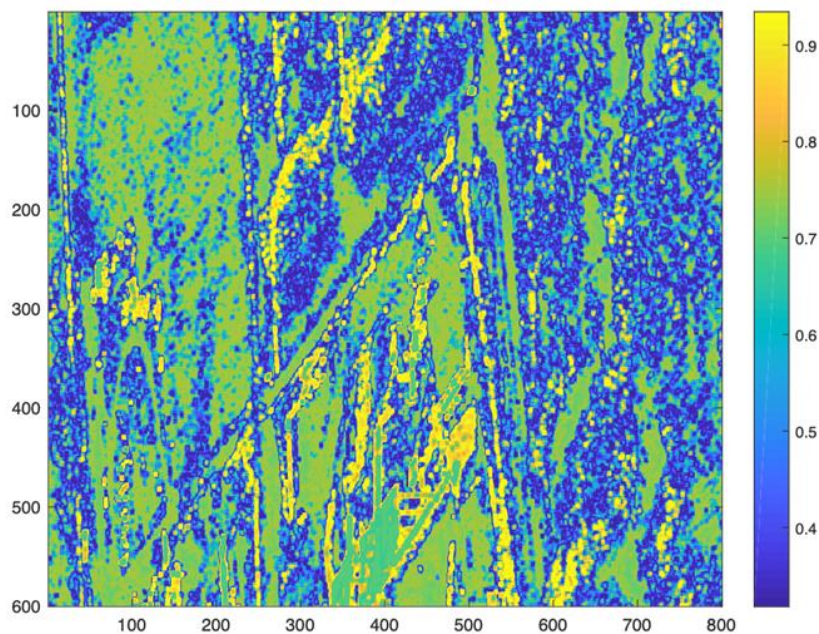
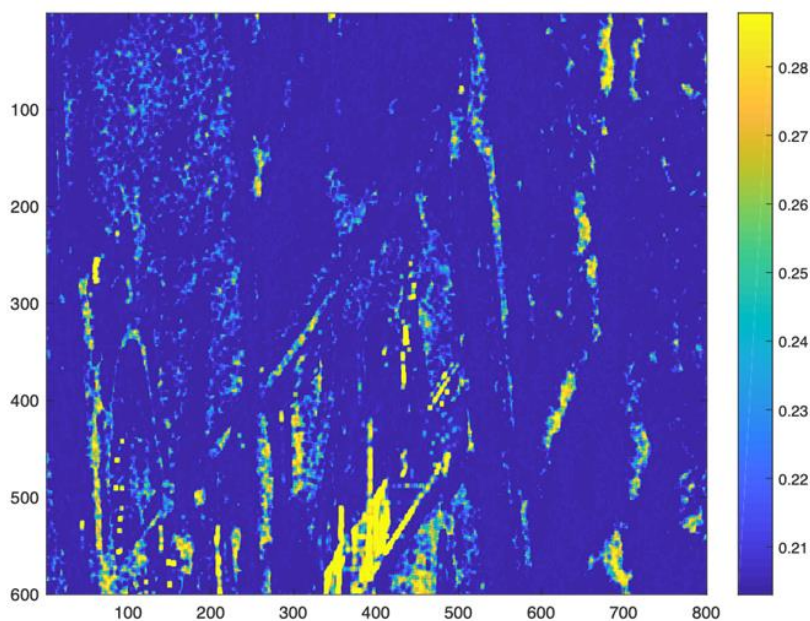
(a) 极化 α 角预测均值(b) 极化 α 角置信水平

图 5-6 X 波段 TerraSAR 数据上模型预测的极化散射角均值及其对应的置信水平

Figure 5-6 Predicted mean value of polarimetric alpha angle and the corresponding certainty level from a TerraSAR-X image

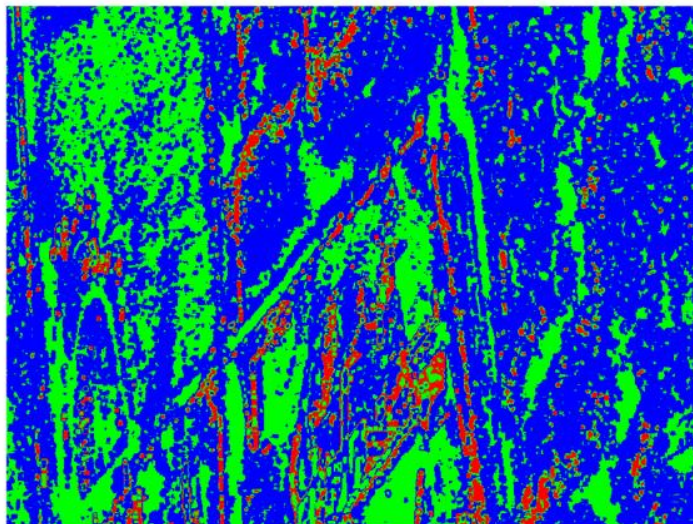


图 5-7 在 TerraSAR-X 影像上的物理散射类型提取结果

Figure 5-7 Identification results of physical radar signatures on a TerraSAR-X image

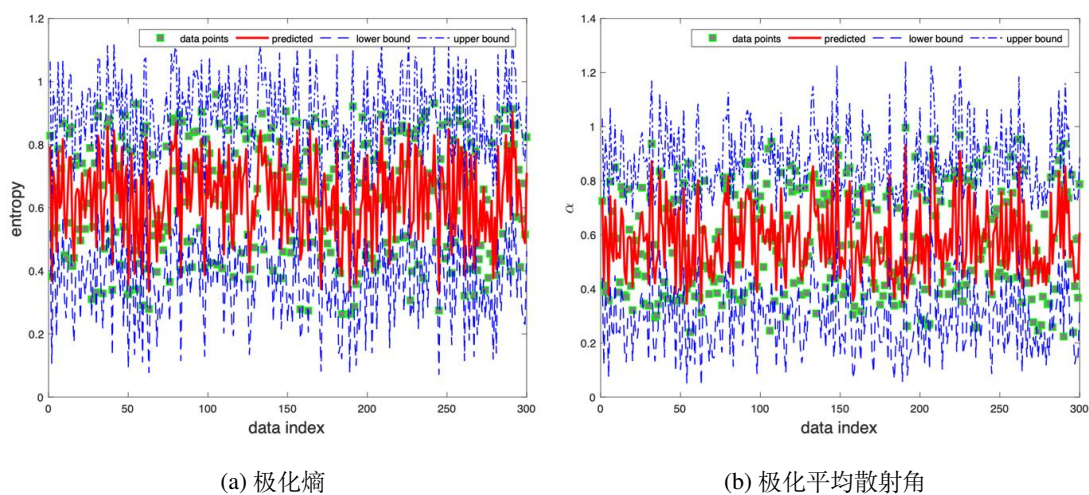


图 5-8 基于 C 波段 Sentinel-1 影像的极化参数训练过程

Figure 5-8 Training process of polarimetric entropy and polarimetric alpha angle on Sentinel-1 data

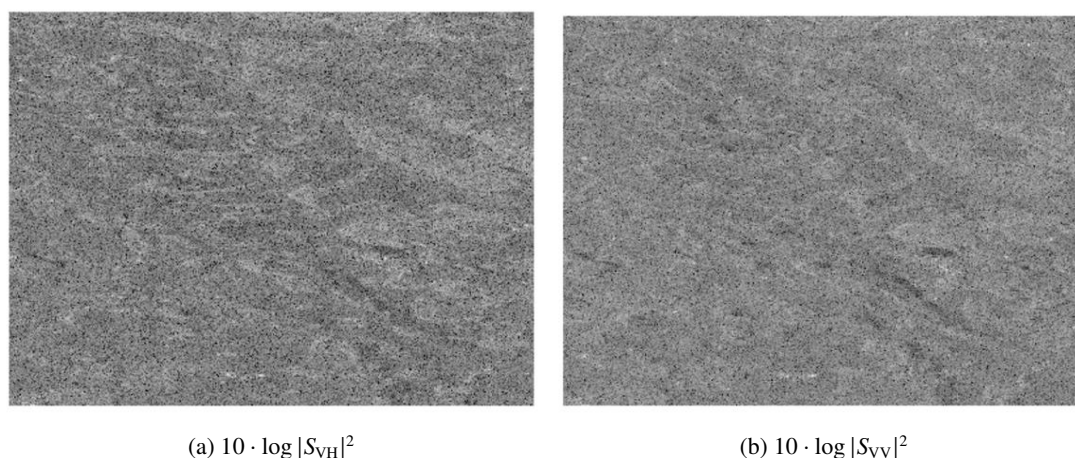
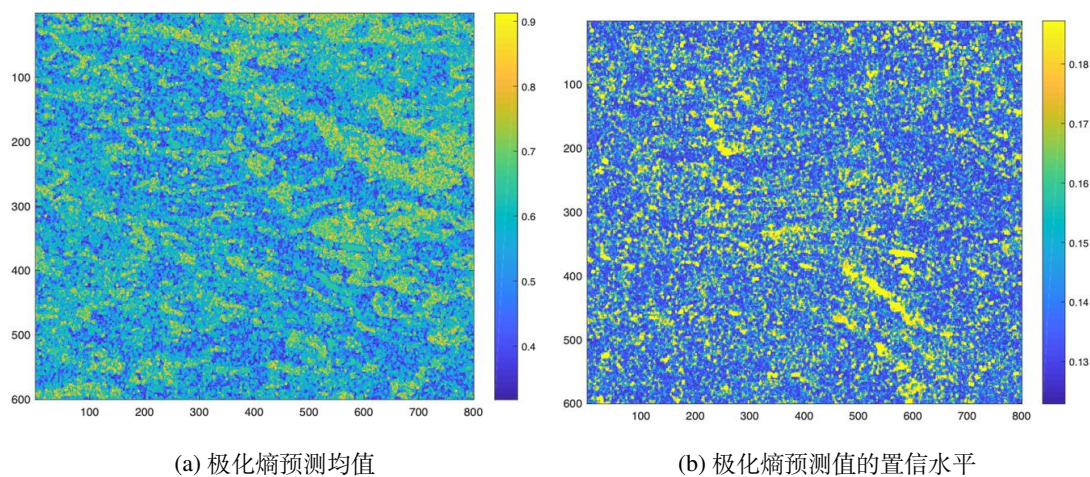


图 5-9 通过对数变换可视化后的 C 波段 Sentinel-1 测试图像

Figure 5-9 C-band Sentinel-1 test image after logarithm transformation

较高，反之表示较低的极化熵预测值。在图 5-10b 中，高亮的取值表示较低的置信水平，反之表示预测值具有较高的置信水平。



(a) 极化熵预测均值

(b) 极化熵预测值的置信水平

图 5-10 基于 C 波段 Sentinel-1 影像预测的极化熵均值和对应的置信水平

Figure 5-10 Predicted mean value of polarimetric entropy from a C-band Sentinel-1 image and the corresponding certainty level

图 5-11 展示了从 C 波段 Sentinel-1 测试图像中所预测的极化平均散射角参数均值和置信水平。其中，图 5-11a 和图 5-11b 分别表示 TerraSAR-X 影像中极化平均散射角的预测均值 $\hat{\alpha}^{\text{sn}}$ 和预测值的置信水平 $\sigma_{\alpha}^{\text{sn}}$ 。高亮的颜色表示高的取值，暗色表示相对较低

的取值。

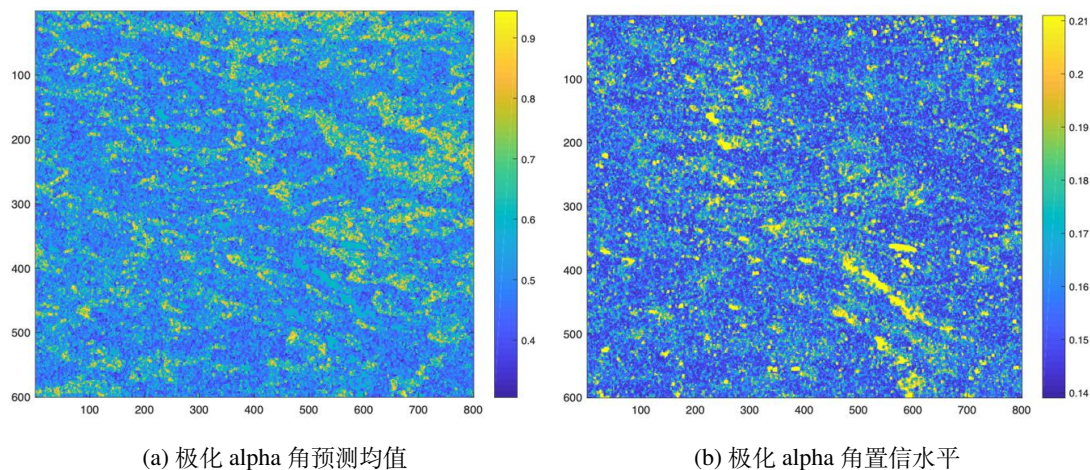


图 5-11 基于 C 波段 Sentinel-1 影像预测的极化散射角均值和对应的置信水平

Figure 5-11 Predicted mean value of polarimetric alpha angle from a C-band Sentinel-1 image and the corresponding certainty level

基于预测的极化熵均值 \hat{H}^{sn} 和极化散射角均值 $\hat{\alpha}^{sn}$ ，本节依据 Cloude 极化目标分解方法和 $H-\alpha$ 极化分割平面得到了 C 波段 Sentinel-1 测试图像的雷达极化物理散射类型。图 5-12 所示为所预测的结果。从图中能够看到所预测的雷达极化物理散射类型能够基本反映图像所对应的地物结构信息。

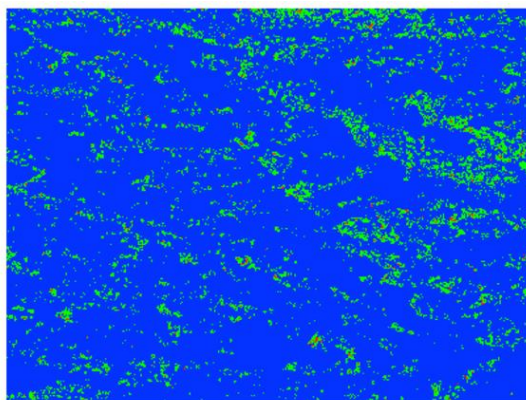


图 5-12 在 C 波段 Sentinel-1 影像上的物理散射类型识别结果

Figure 5-12 Identification of physical radar signatures from a C-band Sentinel-1 image

5.4.4 验证模型的泛化性能

为了验证模型的泛化性能,本文随机选取一景中国上海地区 2018 年 11 月 1 日获取的 C 波段双极化 Sentinel-1 影像,并从多视处理后的影像中截取地物结构较为复杂的 800×1000 像素大小的图像块作为测试图像。所用的 IW 模式 SLC 格式影像经过子带拼接和图像多视等预处理,并提取图像的极化特征。图 5-13 直观地展示了经过对数变换后的可视化图像。其中,图 5-13a 和图 5-13b 分别表示 VH 和 VV 极化方式的图像。这两幅可视化的图像能够比较清晰地给出地物目标的结构信息以及两种不同极化方式下的图像对不同地物目标的后向散射差异。

基于 5.4.3 中 L 波段参考数据 F-SAR 影像和 C 波段源数据 Sentinel-1 影像所训练的高斯过程回归模型和图 5-13 所示的影像,本节预测到该测试图像所对应的极化熵 \hat{H} 和极化平均散射角 $\hat{\alpha}$ 。在此基础上,基于所预测的两个极化散射参数平均值,根据 Cloude 极化目标分解方法和 $H-\alpha$ 极化分割平面从测试数据中推导出图像中具有物理意义的散射类型。在图 5-14 所展示的结果中,红色、绿色和蓝色分别表示多次散射、体散射和表面散射。图中的结果比较清晰地反映了所研究区域的地物结构信息。

5.5 本章小结

本章通过多种探测波段下的 SAR 图像提取雷达极化物理散射类型,从雷达电磁波段的角度分析雷达极化物理散射特性,从而实现面向多种探测波段下 SAR 图像地物散射特性的精细化解译。针对多种探测波段下 SAR 图像的物理散射类型预测问题,首先通过一种基于核函数与平均期望的高斯过程回归的方法预测相关的雷达极化参数(主要包括极化熵和极化散射角及相应的置信度),然后根据 Cloude 极化目标分解和 $H-\alpha$ 极化分割平面从多种探测波段下的 SAR 图像中推导出具有物理意义的地物散射类型。模型训练数据来自于跨波段的源数据和参考数据的重叠区域,并假设单个训练样本的所对应的极化参数服从高斯分布,任意有限个样本的联合分布服从联合高斯分布。在此假设条件下,本章通过 ARD 平方指数核函数使得训练样本和测试样本之间两两关联。高斯过程在预测参数值的同时,能够给出预测值的置信范围,因此,能够实现较为精确的非线性映射。实验过程中所用的参考数据来自德国 Kaufbeuren 地区的 L 波段 F-SAR 影像,源数据来自德国宇航中心的 X 波段 TerraSAR-X 影像和欧空局的 C 波段 Sentinel-1 影像。所训练模型的泛化性能在一幅任意选取的 C 波段 Sentinel-1 影像上得到验证。



(a) $10 \cdot \log |S_{VH}|^2$



(b) $10 \cdot \log |S_{VV}|^2$

图 5-13 C 波段宽幅 Sentinel-1 影像的可视化
Figure 5-13 C-band wide-swath Sentinel-1 test image

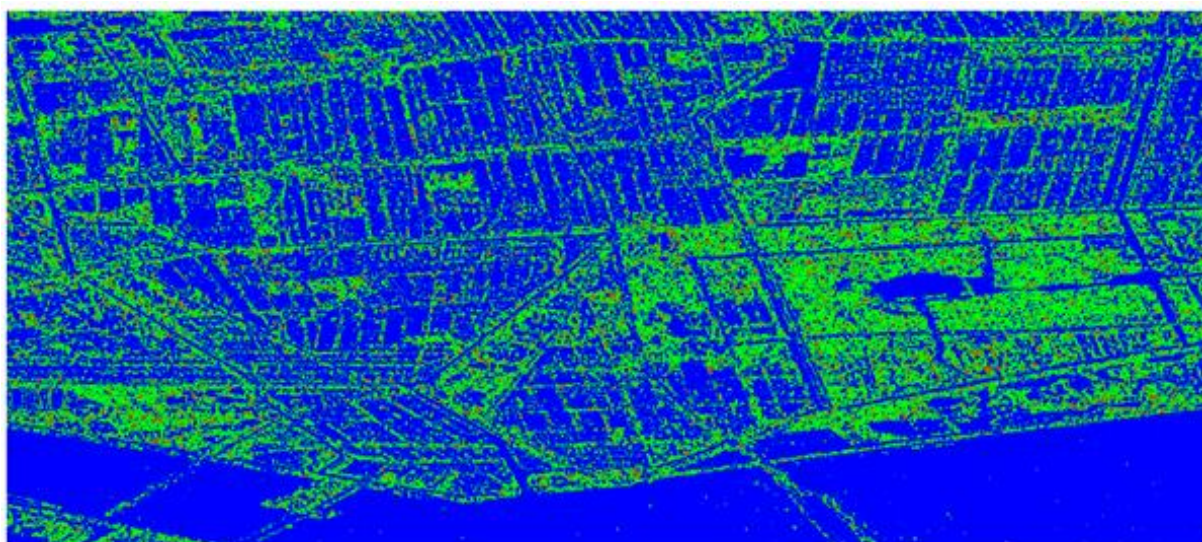


图 5-14 在 Sentinel-1 影像上的物理散射类型识别结果，测试图像尺寸为 800×1000 个像素。红色：多次散射，绿色：体散射，蓝色：表面散射。

Figure 5-14 Identification of physical radar signatures on a Sentinel-1 image. The image size shown in this figure is 800×1000 pixels. R: MS, G: VS, B: SS.

第六章 SAR 图像精细化解译测试库构建与质量评估

6.1 引言

由于对地观测系统，尤其是合成孔径雷达技术，是一个国家国防实力的象征，各个国家相继发射自己的人造地球卫星。SAR 图像已经呈现出海量数据体量、快速增长模式和数据多样化发展的趋势。随着 SAR 成像技术的快速高效发展，SAR 图像领域已经进入大数据时代。部分 SAR 数据也可以公开供科研人员下载与研究使用，这使得针对 SAR 图像解译方法的研究与应用更加便捷高效。大数据发展趋势为面向 SAR 图像的精细化解译带来了更加严峻的挑战。为了高效管理并合理使用 SAR 数据、从中挖掘特定的信息，SAR 图像解译领域亟需发展针对特定成像参数 SAR 数据的测试库，并对其进行质量评估。相应地，测试库发展也是大数据背景下面向 SAR 图像精细化解译的技术基础与数据支撑。

尽管已经有多家单位开始相继建立 SAR 图像样本库，但是到目前为止，公开使用的有标记 SAR 数据非常有限，且其在样本数量和样本的多样性方面并没有非常完善。缺乏对快速增长的、数据量庞大的和数据多样性丰富 SAR 图像的精细化解译研究和合理的利用，因此，建立一个更加完善的供国内外研究者们使用的 SAR 样本库具有一定的实际意义，并能够为多种 SAR 图像解译算法的测试提供数据支撑。

目前，我国在该方面的研究仍处于初级阶段，面向特定成像模式、特定分辨率 SAR 图像的测试库尤其稀缺。这一方面不利于海量多样化数据的有效管理与信息提取；另一方面，SAR 图像测试库的匮乏严重阻碍了后续的数据挖掘、基于深度学习等先进方法的 SAR 解译技术的发展。然而，测试库的构建并非简单的数据收集，加上 SAR 图像不符合人类视觉关注机制的特性造成图像理解难度大，使得 SAR 图像测试库的构建变得异常困难。

本章基于中低分辨率 Sentinel-1 影像，构建了一个面向城市区域场景的 OpenSARUrban 测试库，研究 SAR 图像中测试库的构建方法、测试库的规范化管理，以及基于数据流形可视化和图像分类方法的测试库质量评估。在实施过程中，本章根据 SAR 图像解译系统对测试样本的要求，科学地制定测试库构建路线图、选择样本、对样本进行真值标注、制定科学规范化的管理方案，并采取一系列有效措施对测试库的质量进行评估。

6.2 OpenSARUrban 测试库构建方法

6.2.1 数据收集与预处理

搭载 Sentinel-1 系统的卫星是欧洲空间局哥白尼计划中的对地观测卫星，Sentinel-1 系统共包含两颗卫星，搭载 C 波段合成孔径雷达，连续提供各种天气与气候条件下的对地观测 SAR 影像^[165-167]。尽管这样大规模的 SAR 图像已经能够公开下载¹，但是如何有效利用这些数据并实现 SAR 图像精细化解译需求仍然是研究者们亟需解决的问题。考虑到 SAR 图像解译工具的匮乏、SAR 图像解译技术的自动化智能化程度还不够高，本章构建一个大规模的 SAR 图像精细化解译测试库为各种算法技术的研究提供数据支撑。目前为止，仍然没有针对 Sentinel-1 数据的城市区域土地利用测试库供研究者们使用。为了填补该项空白，本章利用当前公开下载的 Sentinel-1 影像数据，构建了一个面向城市场景的精细化土地利用测试库。

在数据收集阶段，本章首先选取了我国境内的 21 个大型和特大型城市场景的 Sentinel-1 影像作为感兴趣区域，整个数据来源覆盖了我国 17 个省级行政区。表 6-1 列出了 OpenSARUrban 测试库的具体地缘覆盖情况，包括图像覆盖的省级区域和城市区域名称，以及获取图像的具体时间。基本的数据地域分布情况见图 6-1。这幅图直观地展示了数据的地域分布情况，图中灰色和绿色区域分别代表海洋区域和陆地区域，蓝色线条和区域标注了河流和湖泊，红色圆圈表示所标注的城市区域所在地理位置。

测试库构建过程中原始数据来自 Sentinel-1 影像中 Level1 级 IW GRD 模式数据。所选择的 IW GRD 模式通常也被用作默认雷达成像模式。Sentinel-1 产品的数据可随时随地通过欧空局的官网进行免费下载。在测试库构建中，本章采用了原始下载的影像数据和辐射校正后的数据，且二者一一对应。其中，辐射校正数据能够直接反映地物的后向散射特性。图像预处理操作在欧空局发布的 SNAP 应用平台支持下完成。原始数据和辐射校正后的数据均保存为 TIFF 格式供后续使用。

由于 OpenSARUrban 测试库的数据来源于 SAR 影像，其天然的侧视成像机理给人工判读等目视解译带来了很大的难度，使得标注费时费力，测试库构建代价非常昂贵。尽管如此，中低分辨率特性更是加深了目标鉴别的难度。综合考虑以上因素，本文采用光学辅助标注的方法，充分利用了光学影像符合人眼视觉机制的特性、图像的直观性更强、易于人工判读与标注的优势。所采用的光学图像全部来源于 Google Earth，光学图像读取和标注的软件采用北京千帆世景科技有限公司的 91 卫图助手^[168]。该软件支持 20 世纪 30 年代至今的全球数据下载、下载数据无水印、并且支持用户标注以及标注 shapefile 文件的导入导出操作。

¹[https://scihub.copernicus.eu/dhus/\\$#/home](https://scihub.copernicus.eu/dhus/$#/home)

表 6-1 OpenSARUrban 测试库数据来源

Table 6-1 Major Chinese Cities Investigated in This Study

图像覆盖场景 (所覆盖的省级地区)	标注城市名	图像获取时间 (年/月/日)
上海市	上海市	2017/02/03
北京市	北京市	2017/02/01
浙江省	杭州市	2017/03/28
广东省 & 香港	深圳市, 广州市, 香港	2017/02/28
广西省	南宁市	2017/03/10
云南省	昆明市	2017/02/05
四川省	成都市	2017/04/13
重庆市	重庆市	2017/05/02
湖北省	武汉市	2017/01/06
湖南省	长沙市	2017/05/11
江西省	南昌市	2017/03/02
山东省	济南市, 潍坊市	2017/05/15
河北省	石家庄市	2016/11/14
陕西省	西安市	2017/03/05
甘肃省	兰州市	2017/05/15
黑龙江省	哈尔滨市	2016/09/14
福建省	厦门市	2017/05/15
新疆维吾尔自治区	乌鲁木齐市	2017/02/06

6.2.2 精细化类别体系定义

图 6-2所示为 OpenSARUrban 测试库所定义的两级精细化类别标注体系, 图中橘黄色框和浅蓝色框分别描述了一级粗略标注类别和二级精细标注类别。结合实地调研与城市功能评估, 本章所定义的测试库标注体系共分两级, 其中第一级是根据城市功能区将所研究的区域粗略划分为住宅区、商业区、交通枢纽、工业区和其他类别。值得指出的是本章中其他类别指城市绿化带, 这主要是考虑到城市绿化带在城市场景中覆盖面积大的缘故。在这里, 住宅区通常与城市人口变迁有关; 商业区主要与城市经济发展有关; 交通枢纽很大程度上与城市发展的现代化程度有关; 工业区主要与城市产业格局相关联; 而城市绿化主要与人们的健康意识和对生活品质的要求有关。在第二层中, 本章

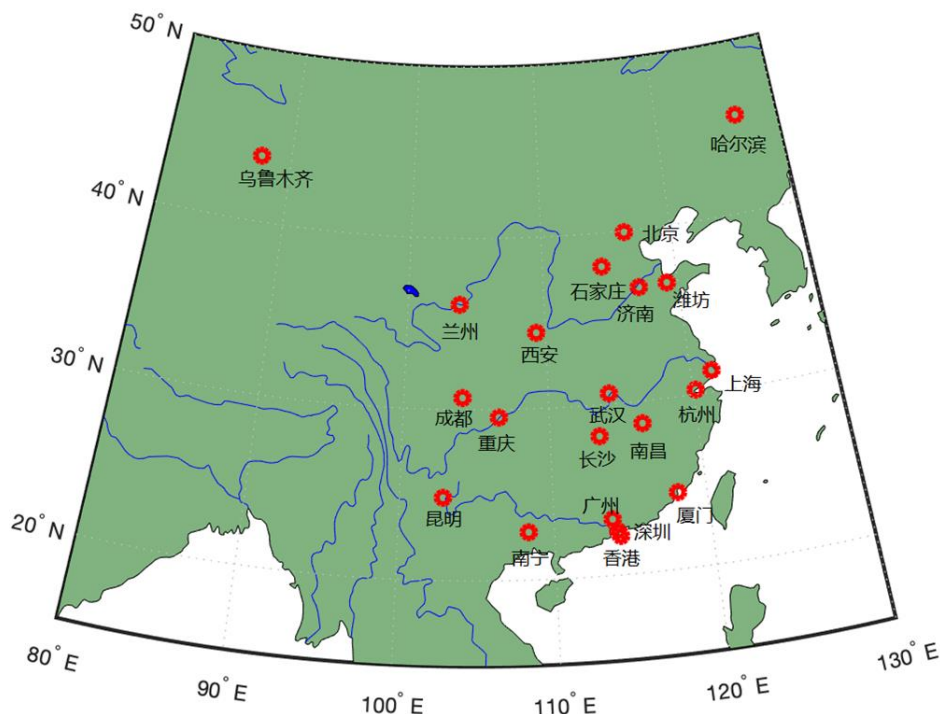


图 6-1 OpenSARUrban 测试库数据源分布情况

Figure 6–1 Data source distributions of the OpenSARUrban dataset

针对第一层中所定义的功能区，对其进行精细化的类别定义。与郊区和农村相比，城市中由于聚集着大量的人口，住宅区便成为一大典型特色。本章根据住宅区域的特点，结合 Sentinel-1 中低分辨率条件下的 SAR 图像特性，将其划分为普通住宅小区、高层住宅区、低矮民房住宅区和独栋住宅区。其中，普通住宅区为我国城市中最常见的居民小区，楼层一般设置为六层、每层有多家住户、楼间距不是非常宽。在高层住宅区中，楼栋高度通常高于六层、每层住户数量较小，并且这类住宅区在设计开发的过程中由于采光需要，往往楼间距非常宽。这种规则的住宅区域在 SAR 图像上往往呈现出比较规则的多层亮线。低矮民房通常是由老旧住宅区域构成，住宅楼区分布非常密集杂乱且无序，在 SAR 图像上也往往呈现出杂乱无规则状态。独栋住宅区主要标注了别墅区，这类建筑区域通常环绕在湖光山色等风景秀丽的地区、周围环境较好、绿化较多。在本数据集中，商业区主要标注了摩天大楼和城市 CBD 等城市中典型的商业区域。交通枢纽囊括了航空枢纽（机场）、铁路枢纽（火车站和高铁站），以及公路枢纽（高速立交）。在该测试库构建中，工业区主要由工业仓储区域组成，该区域一般分布在城市的边缘地带。其他

部分标注了城市绿化带，包括城市中的大型公园、草坪、树林等区域。

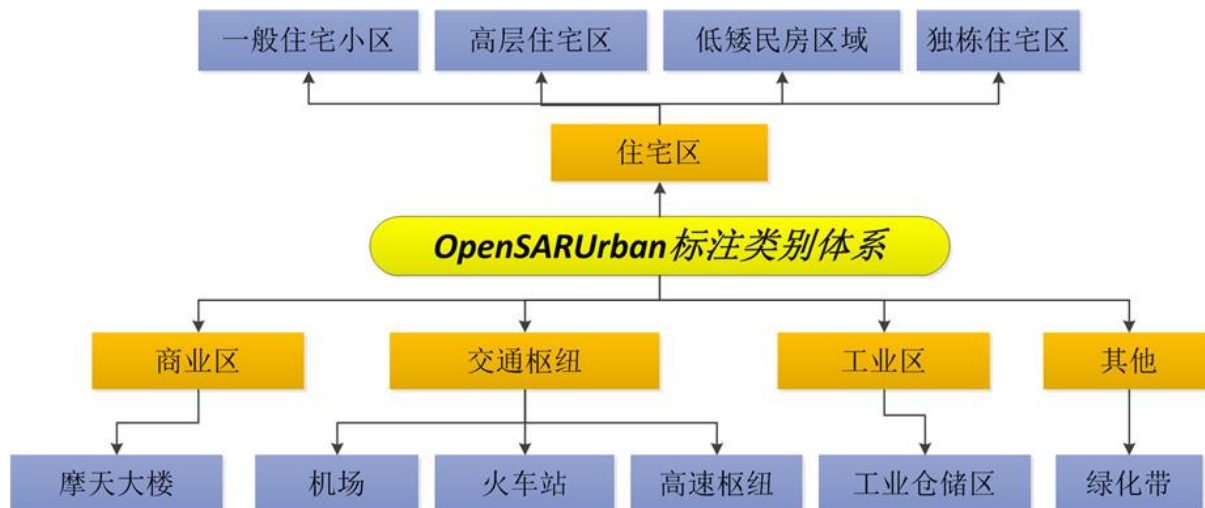


图 6-2 城区两层类别标注体系

Figure 6-2 Two-level hierarchical annotation scheme in urban areas

为了较为直观地展示不同类别之间的数据结构差异，本章对比了该测试库中的 SAR 图像及其对应的光学遥感图像，如图 6-3 所示。

6.2.3 测试库标注方法

OpenSARUrban 测试库构建的具体实施步骤如图 6-4 所示。为了解决中低分辨率 SAR 图像直观标注难度大的问题，本章从 Google Earth 的光学图像中预先标注感兴趣区域的指定类别区域，然后再通过光学和 SAR 图像之间的标注迁移实现最终在 SAR 图像上的标注。在标注过程中，本章首先借助北京千帆世景科技有限公司开发的 91 卫图助手实现特定城区中特定时间段给定类别的光学标注，光学影像中的标注采用多边形的形式记录每个标注多边形顶点位置的地理信息，并以 shapefile 文件的形式保存。然后，在 SNAP 工具包中依次导入所要标注的 Sentinel-1 影像和记录的 shapefile 格式光学标注文件，进而通过该软件的自动定位功能结合人工目视判读将预先产生的光学标注和当前的 SAR 图像进行精确配准，得到 SAR 图像上的标注信息，进而导出该指定标注类别在 SAR 图像上的像素坐标位置。最后，利用 MATLAB 软件，根据最小重叠法原则产生 OpenSARUrban 测试库中的 SAR 图像地物样本切片。总的来说，本章测试库的标注过程是以指定地物类型的多边形地理位置信息为桥梁，实现 SAR 图像和光学图像之间的关联，从而对标注难度非常大的中低分辨率 Sentinel-1 影像进行精细类别标注。下面将详细介绍图 6-4 所示的具体实施步骤。

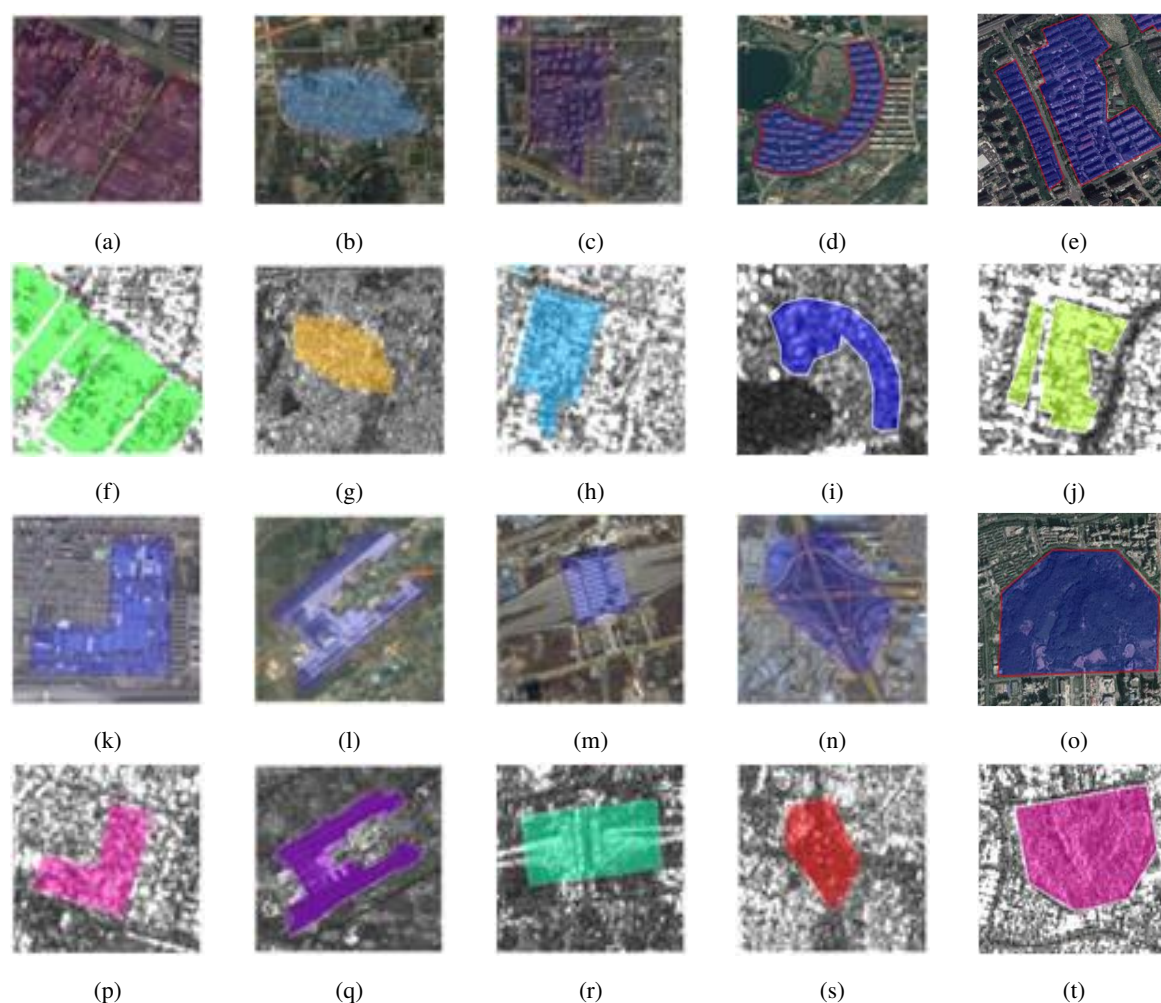


图 6-3 测试库中光学与 SAR 图像样本切片样例对比。第一行和第二行从左至右分别为光学和对应的 SAR 图像摩天大楼、低矮民房、高层建筑区、别墅区和普通住宅区；同样的道理，第三行和第四行分别为光学图像和对应的 SAR 图像，图像类别依次为工业仓储区、航空枢纽、铁路枢纽、高速公路枢纽和城市绿化带区域。

Figure 6-3 Optical examples and their corresponding SAR examples for each category. The first row and the second row show optical and SAR examples of skyscraper, dense and low-rise residential buildings, high-rise buildings, villas, and general residential areas, respectively. The third row and the fourth row show optical and SAR examples of industrial storage areas, airports, railways, highways, and vegetation, respectively. The colored masks are the main land cover locations of the given category.

步骤 1: 影像信息索引。

在海量的光学遥感图像中定位所要标注的数据场景，需要两方面的信息：SAR 影像覆盖区域的地理位置信息（又即感兴趣区域）和影像的成像时间信息，二者均包含在所下载的 Sentinel-1 数据的头文件包中。其中地理位置信息指的是四个角点的精度和纬度，

成像时间信息以“年-月-日”的形式存在。

步骤 2: 定位光学影像。

本章所用的光学影像是 91 卫图助手中的 Google Earth 历史影像，可以通过 91 卫图助手的搜索功能定位到特定场景特定时间段的影像，标注所采用的光学影像分辨率为 0.13 米。值得一提的是 91 卫图支持任意尺度任意幅面数据的高效下载，因此标注不受成像时间段的限制。最后，基于所查询到的位置和时间信息，本章通过 91 卫图助手定位到所要标注的 SAR 影像场景所对应的光学区域。

步骤 3: 创建特定类别光学标注文件夹。

在光学影像中勾勒多边形目标之前，本章需要预先创建一个以所要标注类别命名的文件夹存放光学标注信息。比如当前标注的是高层住宅区，就需要首先创建一个以“High-rise-building”命名的文件夹。然后，在 91 卫图助手中激活所创建的文件夹，并将其作为当前标注文件夹开始标注。这样，标注的信息都会自动存入该激活文件夹中。这个操作能够及时防止标注信息保存在了错误的位置。

步骤 4: 标注光学影像。

对于 20 米分辨率的 SAR 图像来说，Sentinel-1 影像的标注显得非常具有挑战性。在大多数情况下，即使专业的判读专家也无法分辨具体的目标类型。本章重点以街区为单位对城区典型地物目标进行标注。为了克服标注困难的问题，本章充分利用光学影像易辨识的特性，通过光学和 SAR 标注之间的互相关联，并通过坐标转换实现 SAR 影像标注。在光学影像的标注部分，本章借助 91 卫图助手，采用多边形标注的方式，将目标类别以多边形为单位，其顶点的地理位置信息放入 shapefile 文件中。这样，所标注的多边形地理位置信息就能够起到了桥梁的作用，将对应的 SAR 影像和光学影像关联起来。

步骤 5: 保存光学标注。

这一步的目标是保存步骤 4 所标注的 shapefile 文件，将所标注的目标地物类型的地理信息保存在步骤 3 所创建的文件夹中。

步骤 6: 核查光学标注。

接下来需要重复步骤 3~步骤 5，直至类别标注体系中所定义的所有精细类别标注完毕。为了保证标注的可靠性，本章必须在光学标注完成后交互检查核对标注的准确性、所保存的文件信息是否完整、以及存放位置是否准确，从而及时发现问题并校正标注。

步骤 7: 光学和 SAR 图像配准。

本章通过光学标注所保存的 shapefile 地理位置信息并结合人工目视判读，实现影像中光学图像和 SAR 图像的精细配准。在配准之前，本章需要用 SNAP 工具包打开所要标注的 Sentinel-1 影像。然后，以 vector 的形式导入之前保存的 shapefile 光学标注文件。

结合判读专家的人工目视判读，从而实现了光学标注与 Sentinel-1 影像的初步配准。由于受几何畸变的影响，不可避免地存在配准偏移现象。为了解决这个问题，本章也参考了 Google 的光学遥感影像和专家判读知识，在标注层面加入了适当的手动调整，从而实现精确配准。



图 6-4 OpenSARUrban 数据库构建流程图

Figure 6-4 Workflow of compiling the OpenSARUrban dataset

步骤 8: 导出 SAR 影像标注。

这一步是将配准调整后的 SAR 图像标注（像素坐标形式）保存为文本文件，并保存在 SAR 影像标注文件夹下。这里，SAR 影像标注实质上是标注的多边形角点的像素位置信息。

步骤 9: 生成 SAR 图像切片。

具备了标注信息之后，本章以多边形与所要生成的切片的重叠面积不少于 50 个像素为原则，通过滑窗切割法生成最终的测试库样本切片。多边形标注与样本切片的重叠面积计算方法如图 6-5 所示。图中紫色区域和黄色区域分别代表预生成的切片的覆盖范围和所标注的多边形的切割范围，二者的重叠面积是中间重叠区域所占的像素总数。测试库中的样本切片被保存为四种不同的格式，它们分别是：原始 32 位数据、UINT8 格式灰度图、辐射校正后的 SAR 图像和伪彩色图像。所生成的切片尺寸在方位向和距离

向分别为 100×100 像素大小。

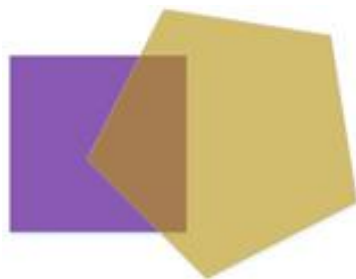


图 6-5 重叠面积计算方法

Figure 6-5 Overlap computation mechanism

6.2.4 测试库组织形式

OpenSARUrban 测试库按照类别标注体系中所定义的精细类别、所保存的数据格式、SAR 图像的极化方式，分类保存在不同的文件夹和子文件夹下。图 6-6 所示为其中一景 Sentinel-1 影像所生成的 OpenSARUrban 数据库的组织形式，其他景影像生成的测试库组织形式基本雷同。每一个位置的样本切片保存为包含 VH 和 VV 两种极化通道的 32 位原始影像数据、UINT8 格式灰度图、辐射校正后的双极化通道数据和伪彩色图。不同格式的数据是为了满足不同的用户需求。目前，测试库中所包含的 10 个城市场景地物类别分别保存在不同的文件夹下。每个切片的命名中包含了其在 Sentinel-1 原始影像中沿方位向和距离向的像素位置、多边形标注索引、和极化方式，从而能够方便特定切片的索引。

与 Sentinel-1 影像的数据格式保持一致，原始数据保存为 32 位数据格式。考虑到 GRD 格式数据的特定形式，原始数据共由两个通道组成，每个通道保存了一种极化方式。基于原始数据，本章做了数据增强以生成 8 位 UINT8 格式的灰度图，该版本数据包括 VH 和 VV 两种极化方式下的 UINT8 格式。图 6-7 展示了每种地物类型下 VH 极化和 VV 极化方式的 UINT8 格式切片。本章通过 SNAP 工具包生成辐射校正图像，并结合标注信息生成了特定地物类型的辐射校正 SAR 切片。辐射校正图像中的数值表征了雷达散射截面积 σ_0 ，该数据格式同样也包含了 VH 和 VV 两种极化通道。

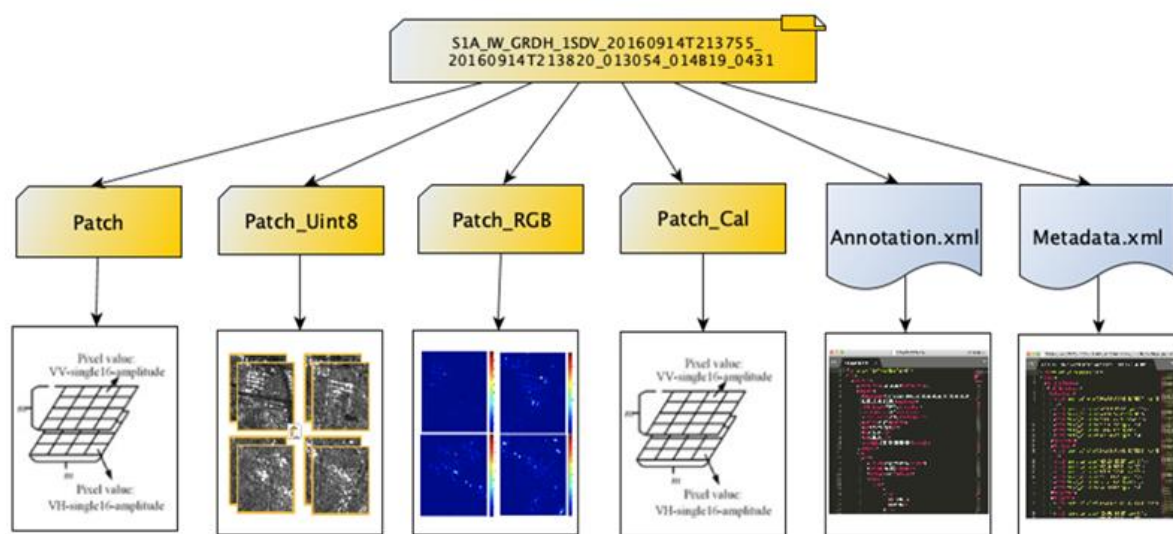


图 6-6 OpenSARUrban 数据库的组织形式

Figure 6-6 Data organization of the OpenSARUrban dataset

6.3 OpenSARUrban 测试库特性分析与规范化管理

6.3.1 测试库特性分析

6.3.1.1 大规模性

OpenSARUrban 测试库提供了一个面向城市场景的中低分辨率测试库。该测试库总共提供了 33,358 个欧空局的 Sentinel-1 卫星获取的 SAR 图像样本切片，共包含 4 种面向不同用户需求的数据格式以及所对应的 2 种不同的极化方式。

6.3.1.2 多样性

数据格式多样性：对于每个样本切片，本测试库共提供了 4 种面向不同用户需求的数据格式，它们分别是：原始 32 位数据、UINT8 格式灰度图、辐射校正数据和伪彩色图像。

地理位置多样性：OpenSARUrban 测试库的数据分布遍及我国境内的 21 个大型或者特大型城市区域，覆盖 17 个省级行政区，其中大多数来源于省会城市场景。图 6-8 所示为所调研的每个城市生成的样本个数。图中不同颜色的柱状图表示每一个城市场景下不同分类类别样本的个数。其中，标记为广东的场景事实上汇总了广州、深圳和香港三个城市的样本，因为这三个城市在地理位置上较近，可以在 Sentinel-1 这样的宽幅 SAR 图像中用同一幅影像呈现出来。

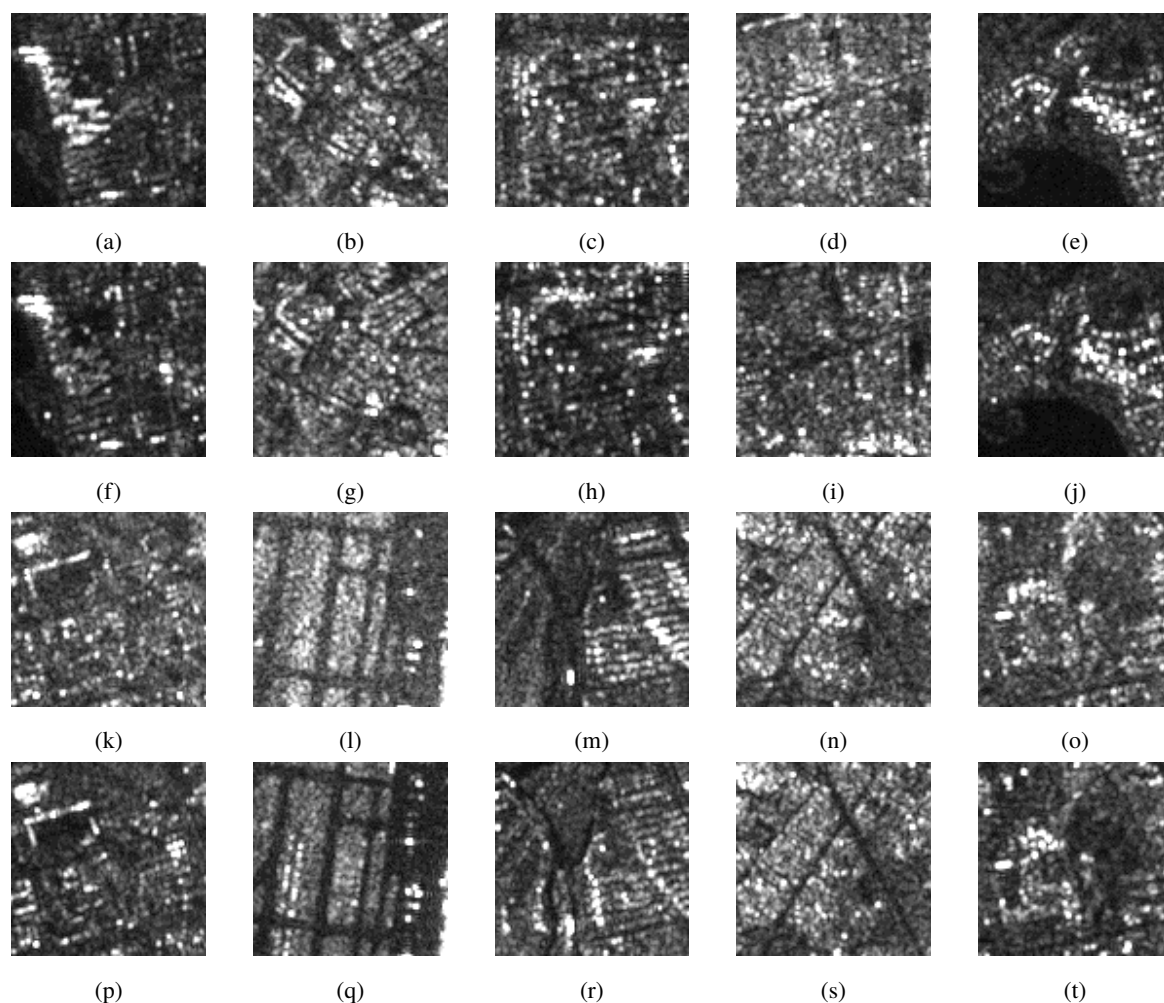


图 6-7 OpenSARUrban 测试库中不同地物类型下的 UINT8 格式切片，包括 VH 极化和 VV 极化两种极化方式，第一行展示了 VH 极化方式测试库样例，从左至右依次是：摩天大楼、普通住宅区、高层住宅区、低矮密集民房区和别墅区；第二行为上一行所对应的 VV 极化方式切片；第三行展示了 VH 极化方式下的测试库样例，从左至右类依次是：航空枢纽、铁路枢纽、高速枢纽、工业仓储区和城市绿化带。

Figure 6-7 Image patches for each category with UINT8 format. The first row shows VH-polarized examples of skyscrapers, general residential areas, high-rise building blocks, dense and low-rise residential areas, and villas; The second row shows the corresponding VV-polarized examples; The third row displays VH-polarized data of airports, railways, highways, industrial storage areas, and vegetated areas; The fourth row exhibits the corresponding VV-polarized patches.

类别多样性：通过实地调研与城市功能评估，本测试库定义了城市场景中 10 类典型地物覆盖区域。在当前中低分辨率 SAR 影像下实现了尽可能精细化的城市功能区类别标注，尤其是在建筑物利用上，首先根据功能区将城市建筑区分为住宅区、工业区

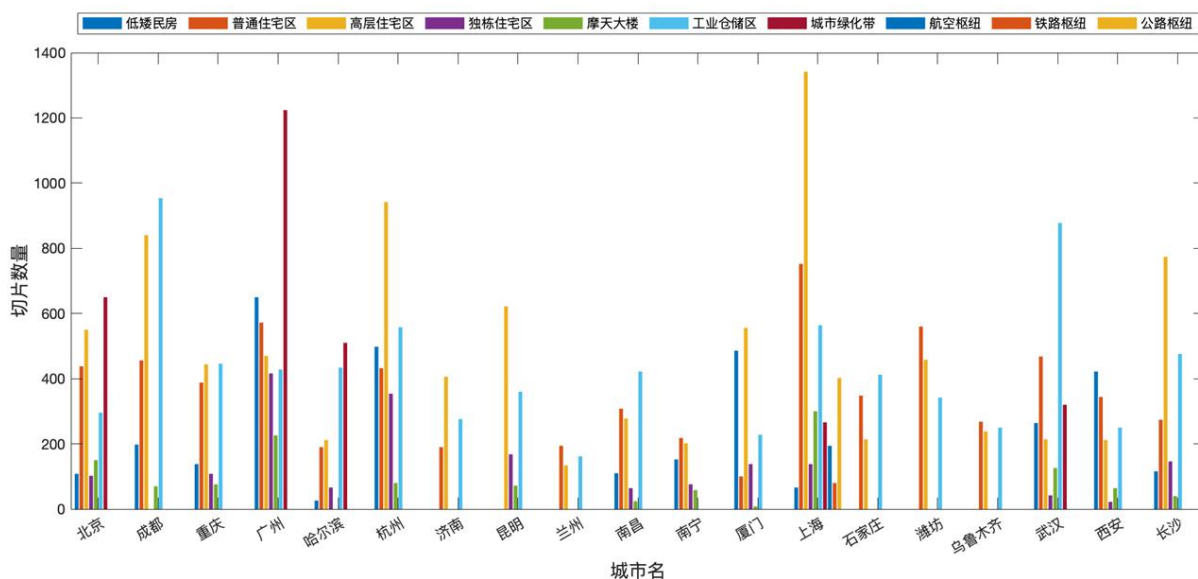


图 6-8 测试库影像地缘分布。不同颜色的柱状图表示不同类别的样本个数。该统计图中的信息按照地理位置分布。

Figure 6-8 Dataset distributions among different cities. Differently colored bars represent the number of image patches from different categories. Image patches are grouped according to the city distributions.

和商业区。在此基础上，本章根据每种功能区在 SAR 图像中的表现特征，将其精细化区分为不同的类别。其中，住宅区分为普通住宅区、高层住宅区、低矮民房和独栋建筑区；工业区和商业区分别以工业仓储区和摩天大楼为代表。这些多样性的类别划分对 Sentinel-1 影像城区的精细化解译具有一定的参考价值，也对城市人口、经济等的变化情况提供了有利条件。

极化方式多样性：为了方便分析雷达极化方式对 Sentinel-1 影像城区解译的影响，OpenSARUrban 测试库包含了 VH 和 VV 两种极化方式的功能区切片，用户能够通过各种先进的算法对影像的极化方式进行定性与定量分析。

6.3.1.3 特定性

OpenSARUrban 测试库针对公开的 Sentinel-1 影像数据提供了一个面向城区解译的大规模测试库，是目前少有的中低分辨率下 C 波段测试库之一。构建该测试库的目的是为城区特定类型的目标解译提供强有力的数据支撑。具体的应用价值包括：1) 研究城市区域中不同地物类型的散射机理和应用潜力；2) 开发更加鲁棒、性能更佳，以及具有针对性的图像分类算法；3) 支持基于内容的城区 SAR 图像检索。

6.3.1.4 可靠性

在测试库构建的过程中，本章通过标注的层层核查与细化来保证测试库的可靠性。为了解决中低分辨率下 SAR 图像目视判读困难的问题，本章通过易于标注的光学图像和 SAR 图像配准的方式及大地提高了标注质量。除此之外，在后续过程中，本章通过数据流形可视化和图像分类性能评估测试库的可靠性。

6.3.1.5 可持续性

光学标注过程中所保存的 shapefile 文件能够用于后续测试库扩充，这将大大降低人力标注的成本，提高标注效率，利于多时相与多模态测试库发展。在未来的研究中，本章将从两方面着手。第一，在现有的类别上继续丰富测试库内容，这需要更新光学标注的 shapefile 文件。第二，做多时相与多模态测试库发展，该测试库将用于变化检测等多时相与多模态分析相关的研究。

6.3.2 测试库规范化管理

OpenSARUrban 测试库最终将按照 Sentinel-1 影像的景名命名，下级子目录分别为数据格式和样本切片所属类别（见图 6-6）。附加一个数据说明文件和详细描述影像参数的头文件，且二者分别保存在 XML 文件中。数据格式总共保存了 4 种类型来满足不同的应用与解译研究需求，它们分别是原始数据保存的 32 位 2 通道矩阵、UINT8 灰度图、辐射校正图像和伪彩色图像。来源于不同影像的不同类别的切片分别保存在不同的子文件夹下。

6.4 测试库质量评估：数据库流形可视化

为了直观地展示本测试库数据的可靠性，本章在数据库质量评估方面，首先采用数据流形可视化的方式展示数据的分布情况。可视化方法采用一种基于快速压缩距离 (Fast Compression Distance, FCD) 和 t-SNE(t-distributed Stochastic Neighbor Embedding) 方法^[169]。其中，文献调研表明 FCD 方法是一种适用于遥感影像度量的方法^[170-172]。OpenSARUrban 测试库可视化的具体实施流程如图 6-9 所示，图中上层和下层分别表示数据状态和处理算法。在具体实施过程中，本章分别针对两种极化方式下的类别样本中随机选取 300 个样本进行可视化，对于样本数量不足的情况，采用给样本加噪的方式对扩充后的样本进行可视化。

OpenSARUrban 测试库流程可视化结果如图 6-10，两幅子图所示结果分别为 VH 极化和 VV 极化两种极化方式数据的流形可视化结果，图中不同颜色表示不同类别的数

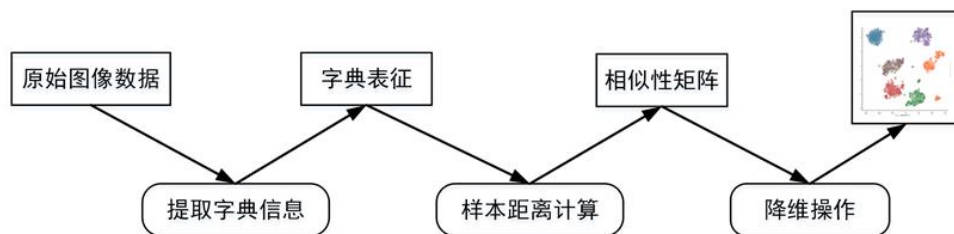


图 6-9 测试库数据流形可视化方法

Figure 6-9 Manifold visualization method for this dataset

据。从可视化的结果可以得出以下结论：

- 从数据库的总体可视化结果对比来看，VH 极化方式的图像区分性更强；
- VH 和 VV 极化方式的图像可视化结果中，交通枢纽（包括航空枢纽、铁路枢纽和公路枢纽）由于其本身结构特征与其他类别差异性较大以及样本扩充带来的影响，导致图像明显区别于其他类别，且 VH 极化方式较 VV 极化方式图像可区分性更强；
- 以城市摩天大楼和 CBD 为代表的城市商业区也能够得到较好的区分，且从流形可视化的角度来说，VH 极化方式图像的区分性更强；
- 涵盖了普通住宅区、高层住宅区、低矮民房和独栋建筑区的城市建筑区域，数据流形基本上聚集在一个聚类区域中，但是其中混杂了少许其他建筑类型的数据；
- 从总体上来看，城市绿化带区域无论是 VH 极化方式还是 VV 极化方式也都能够很好的区分开来，但是从可视化的结果中可以看到，该类别数据与摩天大楼存在小部分交叉现象，原因主要是摩天大楼样本中可能存在部分绿化区域；
- 从可视化的结果也可以看到该测试库在应用上具有一定挑战性。

6.5 测试库质量评估：图像分类

本章通过测试库图像分类的方式定量的对测试库质量进行评估。已有的研究证明 SAR 图像城区场景分类，尤其是中低分辨率的 SAR 图像分类是非常具有挑战性的，研究人员也开发出了多种不同的方法用于 SAR 图像分类识别。本章在测试库评估过程中，采用包括传统基于特征提取与线性 SVM 的方法和基于卷积神经网络的方法，通过图像分类的方式评估本测试库数据的可靠性。

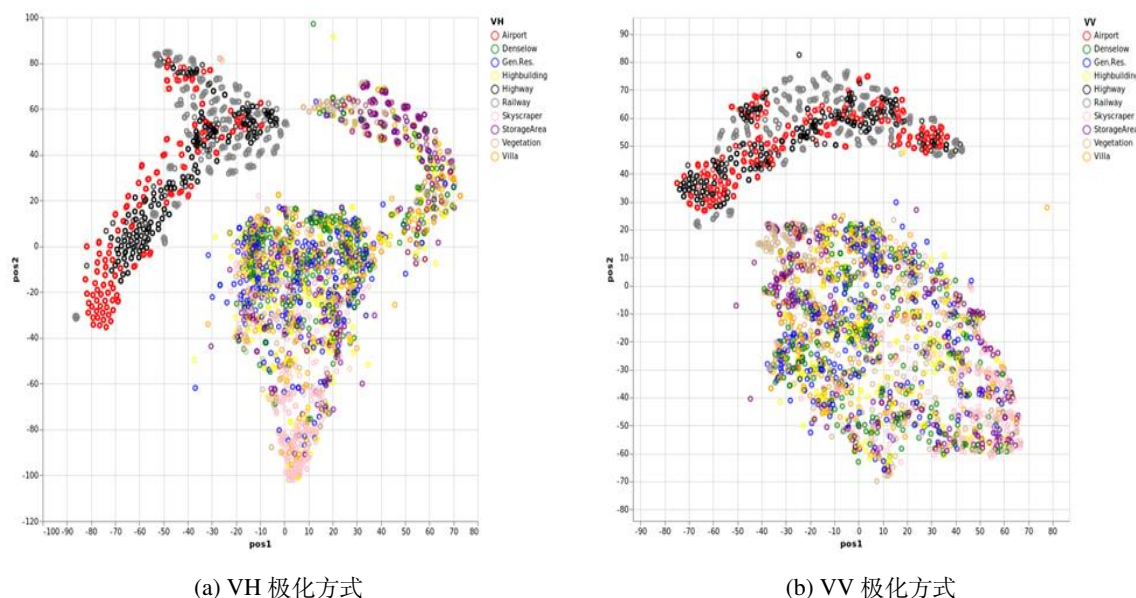


图 6-10 OpenSARUrban 测试库流形可视化结果

Figure 6-10 Urban target visualization of the OpenSARUrban dataset

6.5.1 验证算法与实验设置

评估测试库质量的图像分类算法包括 SAR 图像分类问题中的经典算法和当前主流的卷积神经网络算法。经典 SAR 图像分类算法采用基于特征提取结合线性 SVM 的方法。图像特征提取算法来自局部二值模式方法 (Local Binary Patterns, LBPs)^[173]、LogGabor (Logarithmic Gabor) 方法^[174]、Gabor 方法^[175]、Weber 局部描述子方法 (Weber Local Descriptor, WLD)^[176]、方向梯度直方图 (Histogram of Oriented Gradient, HoG) 算法^[177]和主成分分析算法 (Principle Component Analysis, PCA)^[178]。其中, Gabor 和 LogGabor 特征提取方法的尺度参数和方向参数分别设置为 4 和 6, 这样, 本章计算得到的 Gabor 和 LogGabor 特征是 48 维的列向量。PCA 特征提取方法将每幅图像的原始灰度信息列向量化, 然后通过数据降维的方式得到 30 个维度的主成分向量。为了给出一个公正的评价效果的同时又能够简化问题, 本章的评估算法采用线性 SVM。

鉴于目前主流的深度学习方法, 尤其是在图像分类领域中广泛应用的卷积神经网络方法, 在 SAR 图像分类上的显著优势, 本章也通过一些经典的卷积神经网络学习方法从原始数据中学习图像表示, 实现特征提取与分类器决策端到端一体化框架的方式进行。然后, 通过分类性能, 如分类准确率和混淆矩阵, 综合评价测试库的质量。本章所采用的卷积神经网络算法包括密集连接神经网络 (Densely Connected Convolutional Neural Network, DenseNet)^[179]、50 层深度残差网络 (Residual Neural Network with 50

Layers, ResNet50)^[180]、压缩网络 (SqueezeNet)^[181]、19 层极深卷积神经网络 (Very Deep Convolutional Neural Networks with 19 Layers, VGG19)^[115] 和 AlexNet 网络^[24]。这些网络训练的初始学习率均设置为 0.001, 并且采用 "Poly" 下降策略^[182] 来更新每次迭代中的网络参数直至损失函数收敛。

在该图像分类任务中, 本章从测试库中的每类样本中随机抽取 60% 的体量用于模型训练, 其余样本用于测试。为了解决样本不均衡问题, 本章通过图像加噪的方式扩充模型训练样本。另外, 该图像分类任务对于 VH 极化方式和 VV 极化方式的训练和测试过程完全相同。本章图像分类任务的评价指标采用总体准确率 (Overall Accuracy, OA)^[183, 184] 和混淆矩阵 (Confusion Matrix)^[32, 33]。

6.5.2 测试库分类整体性能评价

基于以上算法设置与评价标准, 对 OpenSARUrban 测试库采用不同分类算法所得的总体准确率如图 6-11 所示。图中紫色和黄色柱状图分别表示 VH 极化方式样本的分类性能和 VV 极化方式样本的性能。从这幅图所展示的效果中可以看到:

- 总的来说, 基于卷积神经网络的方法能够取得比传统方法更好的效果 (LBP 方法除外);
- 对于 VH 和 VV 极化方式的测试库来说, 通过 VGG19 网络的整体准确率均取得了所有评价方法中最好的效果, 且二者的整体准确率分别可以达到 89.49% 和 89.53%;
- 对比所有的基于传统方法的图像分类性能, 不难发现基于 LBP 特征的图像分类方法在 VH 和 VV 极化方式的测试库上均能够取得最好的分类性能 (整体准确率分别是 71.06% 和 70.82%), 并且能够超越某些深度学习方法。

在实际情况中, 分类性能取决于多种因素, 如训练样本质量、特征选择等。然而, 通常 VH 极化方式的人造目标比 VV 极化方式的目标能够取得更好的性能。这可以从以下几点加以解释: 1) 如果从雷达散射理论的角度来理解, 本章研究内容所涉及到的类别大多来自城市中的人造目标 (除城市绿化带之外), 而对人造目标来说, VH 极化方式的视觉效果通常比 VV 极化方式的更好^[185]; 2) 从视觉效果上来说, VH 极化方式的样本能够展现出更好的视觉效果, 从图 6-7 所展示的 VH 和 VV 极化方式的 UINT8 格式对比样例中可以看出; 3) 从测试库流形可视化的结果来看, VH 极化方式的类别之间区分度更强。这可以从图 6-10a 和图 6-10a 流形可视化的对比结果看出。

为了更加客观实际地展示 OpenSARUrban 测试库的分类性能, 本章也提供了基于多种经典卷积神经网络方法和传统方法中取得最好性能的混淆矩阵。在这两类方法中取得最佳图像分类整体准确率的方法分别是 VGG19 网络和基于 LBP 特征提取的 SVM 方法。

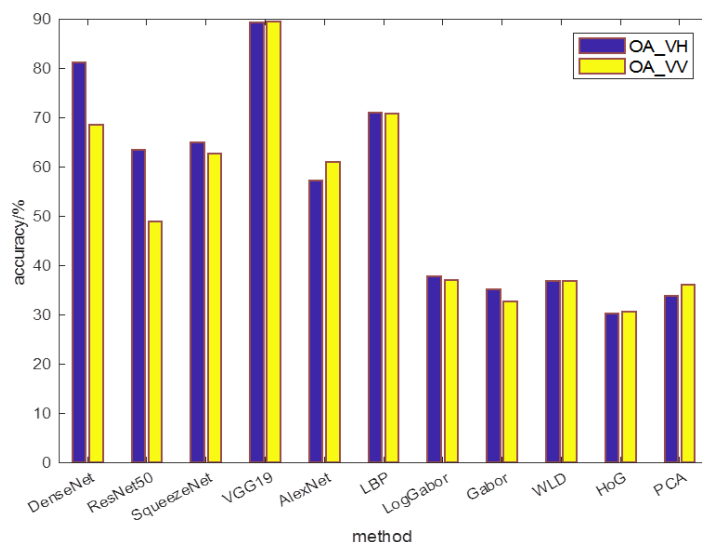


图 6-11 对整个测试库采用不同评价方法所得的总体准确率

Figure 6-11 Overall classification accuracy of the whole dataset with different method

图6-12a和图6-12b分别展示了基于 VGG19 网络的 VH 和 VV 极化方式测试库的测试混淆矩阵。通过观察并对比基于 VGG19 网络所得到的这两个混淆矩阵，容易看出尽管测试库数据在流形结构上存在一定的混淆性，但是通过卷积神经网络的方法仍然能够取得较好的分类效果。但是，本章能够发现针对特定建筑物类型：普通住宅区、低矮民房和工业仓储区的分类效果依然存在较大的提升空间，这是由中低分辨率下 SAR 图像的识别性能降低所造成的。

基于人工设计特征 LBP 和线性 SVM 方法的测试库图像分类结果中的混淆矩阵如图 6-13 所示。其中，图 6-13a 和图 6-13b 分别为 VH 和 VV 极化方式测试库 UINT8 灰度图像的混淆矩阵。从这两个混淆矩阵能够得出以下结论：1) 除了三种典型建筑区域（普通住宅区、低矮民房和工业仓储区）之外，测试库中的大多数类别均能够通过 VGG19 网络模型取得较好的分类效果；2) 这三个典型的建筑区图像极易相互混淆，验证了中低分辨率 SAR 图像的区分性不强；3) 基于传统 LBP 特征和线性 SVM 的分类方法在 VH 极化数据和 VV 极化数据上取得的分类效果没有明显差异。以上这些结论，一方面验证了包括 LBP 特征在内的传统手工设计特征在中低分辨率 SAR 图像表征上的局限性；另一方面，能够看到三类交通枢纽由于其本身结构形态上的差异性，采用 LBP 方法依然能够取得较好的区分效果，这也说明了传统方法在某些应用方面具有其自身的优势。

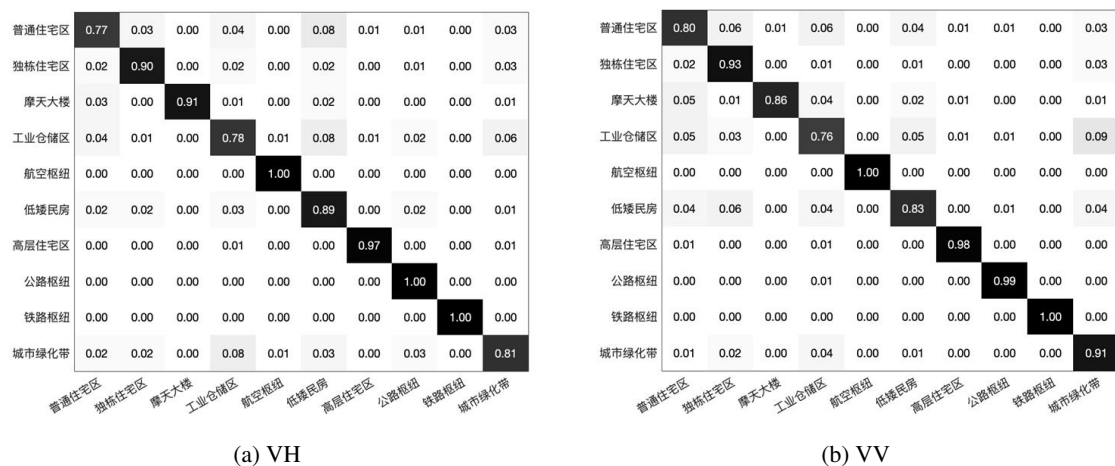


图 6-12 基于 VGG19 网络的不同极化方式测试库分类混淆矩阵

Figure 6-12 Confusion matrices by VGG19 when classifying the whole dataset

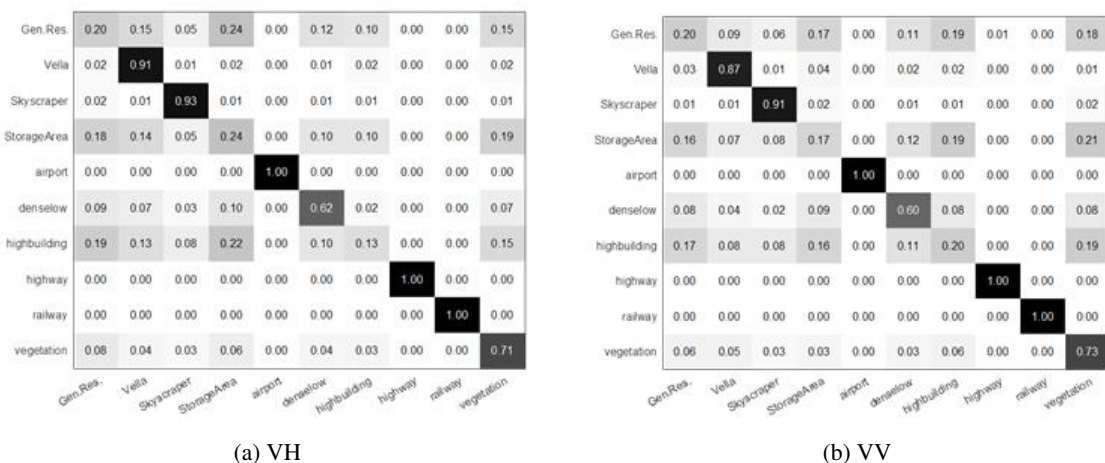


图 6-13 基于 LBP 特征的不同极化方式测试库分类混淆矩阵

Figure 6-13 Confusion matrices by LBP when classifying the whole dataset

6.6 本章小结

本章的主要贡献在于针对中低分辨率的 Sentinel-1 影像中的城区土地覆盖类型, 定义了一个面向城区土地利用类型的精细化两级分类体系, 提供了一个包含 33,358 个 SAR 图像、沿方位向和距离向样本尺寸为 100×100 个像素大小的 SAR 图像测试库。在该测试库中, 每个 SAR 图像切片保存了面向不同应用需求的 4 种数据格式 (分别是 32 位原始数据、辐射校正数据、UINT8 格式灰度图和伪彩色图像)、2 种不同的极化方式 (分别是 VH 和 VV 极化方式)。概括地讲, 测试库具有 5 种典型特性: 大规模性、多样性、特定性、可靠性和可持续发展性。这些性能保证了测试库能够面向三种 SAR 图像解译中的应用方向提供数据支撑, 这几种应用问题包括: 1) 支持中低分辨率下 SAR 图像城市区域特定类型的数据特性研究; 2) 支持发展针对该类数据的多种高级算法研究与测试; 3) 支持基于内容的 SAR 图像检索研究。

然而, 由于测试库发展耗时费力, 除了相关的算法研究之外, 该测试库在未来的研究中, 仍然可以针对以下几个方面继续细化丰富: 1) 由于数据标注过程中保存了基于光学标注的 shapefile 文件, 可以在此基础上拓展多时相城区样本测试库; 2) 同样基于已获取的 shapefile 文件, 可以在光学遥感等其他源的数据上拓展, 实现多模态城区样本测试库构建。这些测试库都将为 SAR 图像解译中面向不同研究方向的应用问题与方法研究提供数据支撑, 具有一定的实用价值。

除此之外, 考虑到本章所定义的精细化层级分类体系结构中, 针对不同的城市功能区与特定城市发展指标密切相关, 如城市住宅区与人口密度相关、商业区与城市发展的现代化程度相关、工业仓储区与城市产业结构相关等, 在未来的研究中, 将基于 SAR 图像定性或定量分析城市规划与发展态势, 并对城市规划的合理性进行评估。

全文总结

SAR 图像解译在军事和民用领域均具有广泛的研究与应用。近年来, 基于卷积神经网络等机器学习方法的 SAR 图像解译更是取得了突破性的研究进展。本文在 SAR 成像技术快速发展的背景下, 重点研究基于卷积神经网络和高斯过程回归等机器学习方法的 SAR 图像目标和地物精细化解译。研究过程中所采用的数据主要来自欧空局的 Sentinel-1 影像、德国宇航中心的 TerraSAR-X 影像和 F-SAR 影像、以及我国自主研制的高分三号影像。总的来说, 本文的主要贡献可以归纳为以下几点:

1. 本文首先通过改进的多尺度卷积神经网络同时实现了 SAR 图像中远海海域和近港/岸等目标密集排布区域的舰船目标高精度检测。在此基础上, 为了解决检测框精确度不足、近港/岸区域漏检严重, 以及少量陆地虚警的问题, 本文提出了一种结合级联卷积神经网络和脉冲余弦变换的方法, 用于实现 SAR 图像中舰船目标的精细化提取。前者较为充分地挖掘了 SAR 图像中的空间域信息, 后者主要在局部自适应区域内通过图像的频域处理实施精细化提取, 最终在 DEM 数据的辅助下实现了面向 SAR 图像舰船目标检测的精细化解译。
2. 为了通过单极化和双极化 SAR 图像提取具有物理意义的散射类型, 并探索多种不同极化方式及其组合下 SAR 图像对物理散射特性的可解释性能, 本文提出了一种基于复数域卷积神经网络的方法, 网络优化的损失函数结合了复数域基础损失项和对比约束项。该方法能够直接从原始复 Sinclair 散射矩阵中学习具备雷达极化物理特性表征的复数域模型, 从而实现多种极化方式下 SAR 图像物理散射类型提取的精细化解译。该方法的有效性和泛化性能在德国宇航中心的多景 L 波段 F-SAR 数据上得到了验证。
3. 为了实现多种探测波段下 SAR 图像地物散射特性的精细化解译, 本文首先通过一种基于核函数与平均期望的高斯过程回归方法预测极化熵和极化散射角度的均值, 以及相应参数的置信范围。然后, 在此基础上通过 Cloude 的 $H - \alpha$ 划分准则提取到相应探测波段 SAR 图像中具有物理意义的散射类型。高斯过程回归模型在预测极化参数均值的同时也能够给出预测值的置信度, 从而能够以较高的精确度实现多种探测波段数据之间的极化参数非线性映射。核机制与平均期望操作能够缓和多波段数据之间几何与辐射畸变不一致带来的影响。该方法的有效性和泛化性能在实测 L 波段 F-SAR 数据、X 波段 TerraSAR-X 数据、以及 C 波段 Sentinel-1 数据上得到了验证。

4. 在海量 SAR 图像数据快速发展的背景下, 本文针对中低分辨率 SAR 图像地物测试库匮乏的问题, 在预定义的精细化两级城区类别标注体系基础上, 构建了一个包含 33,358 个样本的 SAR 图像地物精细化解译测试库 OpenSARUrban。测试库数据来源于 C 波段 IW GRD 模式的 Sentinel-1 影像。测试库中每个切片对应 VH 和 VV 两种极化方式; 并保存了原始 32 位数据、UINT8 格式数据、辐射校正图像和伪彩色图共四种不同的数据格式。测试库具有的大规模特性、多样性、特定性、可靠性、可持续发展特性保证了多种面向 SAR 图像精细化解译算法的实施。此外, 本文从数据流形可视化和图像分类的角度分别对测试库进行质量评估。

本文的研究工作主要基于卷积神经网络和高斯过程回归方法展开, 重点研究了大数据背景下 SAR 图像精细化解译当中的几个问题: 面向 SAR 图像舰船检测的目标精细化解译、多种极化方式下 SAR 图像物理散射类型提取、多种探测波段下 SAR 图像物理散射类型提取, 以及 SAR 图像地物精细化解译测试库发展与质量评估。鉴于所用方法的先进性和海量 SAR 数据的大爆发性, 在未来的研究工作中, SAR 图像的精细化解译可以从以下几方面继续展开研究与探索:

1. SAR 图像中的幅度和强度数据损失了原始的复数域数据结构, 可以重点开展面向 SAR 图像解译的复数域卷积神经网络方法研究。
2. 当前基于卷积神经网络方法的 SAR 图像解译多重视挖掘图像的空间域特性, 缺乏对频域信息的有效利用, 未来可以研究面向 SAR 图像解译的时频结合深度卷积神经网络方法。
3. 在大规模 SAR 图像测试库发展问题上, 当前的测试库数据的多样性 (包括地缘多样性、极化方式、分辨率等等) 仍然不够, 数据不均衡问题突出, 类别有限, 后续可以继续针对特定问题构建相应的测试库; 此外, 也可以在本文基础上发展多模态、多时相测试库, 从而满足相应的研究需求。
4. 研究结合专家知识的智能化 SAR 图像解译系统。当前基于卷积神经网络等机器学习方法的 SAR 图像解译大多是直接从数据中学习模型, 在未来的研究中, 可以通过引入专家先验知识和脑电信号等, 同时从数据和知识中学习 SAR 图像解译模型, 从而提高 SAR 图像的智能化解译水平。

参考文献

- [1] 张国敏. 复杂场景遥感图像目标检测方法研究[D]. 国防科学技术大学, 2010.
- [2] 王琼华. 多分辨率遥感图像复合分类方法研究[D]. 清华大学, 2011.
- [3] JENSEN J R, et al. Introductory digital image processing: a remote sensing perspective.[M]. Prentice-Hall Inc., 1996.
- [4] LILLESAND T, KIEFER R W, CHIPMAN J. Remote sensing and image interpretation[M]. John Wiley & Sons, 2015.
- [5] 吴一戎. 合成孔径雷达散射波干扰研究[J]. 电子学报, 2002, 30(12): 1882-1884.
- [6] 魏钟铨, 等. 合成孔径雷达卫星[M]. 科学出版社, 2001.
- [7] CURLANDER J C, MCDONOUGH R N. Synthetic aperture radar[M]. Wiley, New York, 1991.
- [8] CUMMING I G, WONG F H. Digital processing of synthetic aperture radar data[J]. Artech house, 2005, 1(3).
- [9] 王超, 张波, 温晓阳, 等. 基于雷达散射特性的高分辨率 SAR 图像自动目标识别[J]. 电波科学学报, 2004, 19(4): 422-426.
- [10] 占必超, 吴一全, 纪守新, 等. 基于平稳小波变换和 Retinex 的红外图像增强方法[J]. 光学学报, 2010, 30(10): 2788-2793.
- [11] 朱博, 王新鸿, 唐伶俐, 等. 光学遥感图像信噪比评估方法研究进展[J]. 遥感技术与应用, 2010, 25(2): 303-309.
- [12] 叶滢钰. 基于稀疏表示和随机观测的 SAR 图像分割. 西安电子科技大学, 2014.
- [13] 焦李成, 王爽, 侯彪. SAR 图像理解与解译研究进展[J]. 电子学报, 2005, 33(B12): 2423-2434.
- [14] BLACKNELL D, ARINI N S, MCCONNELL I. SAR image understanding using contextual information[C]//SAR Image Analysis, Modeling, and Techniques IV:vol. 4543. International Society for Optics, 2002: 73-84.

- [15] SOERGEL U, SCHULZ K, THOENNESSEN U, et al. Integration of 3D data in SAR mission planning and image interpretation in urban areas[J]. *Information Fusion*, 2005, 6(4): 301-310.
- [16] HELLMANN M, KROGAGER E. Comparison of decompositions for pol-SAR image interpretation[C]//IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No. 00CH37120):vol. 3. 2000: 1313-1315.
- [17] CRISP D J. A ship detection system for RADARSAT-2 dual-pol multi-look imagery implemented in the ADSS[C]//2013 International Conference on Radar. 2013: 318-323.
- [18] BRUSCH S, LEHNER S, FRITZ T, et al. Ship surveillance with TerraSAR-X[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2010, 49(3): 1092-1103.
- [19] ESCH T, TAUBENBÖCK H, ROTH A, et al. TanDEM-X mission-new perspectives for the inventory and monitoring of global settlement patterns[J]. *Journal of Applied Remote Sensing*, 2012, 6(1): 061702.
- [20] COVELLO F, BATTAZZA F, COLETTA A, et al. COSMO-SkyMed an existing opportunity for observing the Earth[J]. *Journal of Geodynamics*, 2010, 49(3-4): 171-180.
- [21] MALENOVSK Z, ROTT H, CIHLAR J, et al. Sentinels for science: Potential of Sentinel-1,-2, and-3 missions for scientific observations of ocean, cryosphere, and land[J]. *Remote Sensing of Environment*, 2012, 120: 91-101.
- [22] MARINO A, SUGIMOTO M, OUCHI K, et al. Validating a notch filter for detection of targets at sea with ALOS-PALSAR data: Tokyo Bay[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2013, 7(12): 4907-4918.
- [23] LEE S R. Overview of KOMPSAT-5 program, mission, and system[C]//IEEE International Geoscience and Remote Sensing Symposium. 2010: 797-800.
- [24] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]//Advances in neural information processing systems. 2012: 1097-1105.
- [25] 丁鹏. 基于深度卷积神经网络的光学遥感目标检测技术研究[D]. 中国科学院大学(中国科学院长春光学精密机械与物理研究所), 2019.

- [26] LI H, LIN Z, SHEN X, et al. A convolutional neural network cascade for face detection[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015: 5325-5334.
- [27] RANJAN R, PATEL V M, CHELLAPPA R. HyperFace: A Deep Multi-Task Learning Framework for Face Detection, Landmark Localization, Pose Estimation, and Gender Recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 41(1): 121-135.
- [28] MA X, DAI Z, HE Z, et al. Learning Traffic as Images: A Deep Convolutional Neural Network for Large-Scale Transportation Network Speed Prediction[J]. Sensors, 2017, 17(4): 818.
- [29] DABIRI S, HEASLIP K. Inferring transportation modes from GPS trajectories using a convolutional neural network[J]. Transportation Research Part C-emerging Technologies, 2018, 86: 360-371.
- [30] DUMITRU C O, CUI S, SCHWARZ G, et al. Information content of very-high-resolution SAR images: Semantics, geospatial context, and ontologies[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2014, 8(4): 1635-1650.
- [31] YANG Y, NEWSAM S. Bag-of-visual-words and spatial extensions for land-use classification[C]//Proceedings of the 18th SIGSPATIAL International Conference on Advanced Geography Information System (ACM GIS). San Jose, CA, USA, 2010: 270-279.
- [32] XIA G S, HU J, HU F, et al. AID: A benchmark data set for performance evaluation of aerial scene classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(7): 3965-3981.
- [33] JIN P, XIA G S, HU F, et al. AID++: An Updated Version of AID on Scene Classification[C]//Proceedings on IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Valencia, Spain, 2018: 4721-4724.
- [34] XIA G S, BAI X, DING J, et al. DOTA: A large-scale dataset for object detection in aerial images[C]//Proceedings on IEEE Conference of Computer Vision and Pattern Recognition. Salt Lake City, Utah, 2018: 3974-3983.

- [35] KEYDEL E R, LEE S W, MOORE J T. MSTAR extended operating conditions: A tutorial[C]//Algorithms for Synthetic Aperture Radar Imagery III:vol. 2757. 1996: 228-243.
- [36] CHEN S, WANG H, XU F, et al. Target classification using the deep convolutional networks for SAR images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(8): 4806-4817.
- [37] PEI J, HUANG Y, HUO W, et al. SAR automatic target recognition based on multiview deep learning framework[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 56(4): 2196-2210.
- [38] YUAN X, TANG T, XIANG D, et al. Target recognition in SAR imagery based on local gradient ratio pattern[J]. International Journal of Remote Sensing, 2014, 35(3): 857-870.
- [39] HOFFMANN K, FISCHER P. DOSAR : a multifrequency polarimetric and interferometric airborne SAR-system[C]//IEEE International Geoscience and Remote Sensing Symposium:vol. 3. 2002: 1708-1710.
- [40] DUMITRU C O, SCHWARZ G, DATCU M. SAR Image Land Cover Datasets for Classification Benchmarking of Temporal Changes[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(5): 1571-1592.
- [41] HUANG L, LIU B, LI B, et al. OpenSARShip: A dataset dedicated to Sentinel-1 ship interpretation[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(1): 195-208.
- [42] SCHMITT M, HUGHES L H, ZHU X X. The SEN1-2 Dataset for Deep Learning in SAR-Optical Data Fusion[J/OL]., 2018. <https://arxiv.org/abs/1807.01569>.
- [43] ZHAO J, GUO W, CUI S, et al. Convolutional neural network for SAR image classification at patch level[C]//IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 2016: 945-948.
- [44] CHEN S, WANG H. SAR target recognition based on deep learning[C]//International Conference on Data Science and Advanced Analytics (DSAA). 2014: 541-547.
- [45] LI X, LI C, WANG P, et al. SAR ATR based on dividing CNN into CAE and SNN[C]//IEEE 5th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR). 2015: 676-679.

- [46] VINCENT P, LAROCHELLE H, LAJOIE I, et al. Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion[J]. Journal of Machine Learning Research, 2010, 11(110): 3371-3408.
- [47] DING J, CHEN B, LIU H, et al. Convolutional neural network with data augmentation for SAR target recognition[J]. IEEE Geoscience and remote sensing letters, 2016, 13(3): 364-368.
- [48] ZHANG Z, WANG H, XU F, et al. Complex-valued convolutional neural network and its application in polarimetric SAR image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(12): 7177-7188.
- [49] STEENSON B O. Detection performance of a mean-level threshold[J]. IEEE transactions on Aerospace and Electronic Systems, 1968(4): 529-534.
- [50] GANDHI P P, KASSAM S A. Analysis of CFAR processors in nonhomogeneous background[J]. IEEE Transactions on Aerospace and Electronic systems, 1988, 24(4): 427-445.
- [51] NOVAK L, HESSE S. On the performance of order-statistics CFAR detectors[C]// [1991] Conference Record of the Twenty-Fifth Asilomar Conference on Signals, Systems & Computers. 1991: 835-840.
- [52] SMITH M E, VARSHNEY P K. VI-CFAR: A novel CFAR algorithm based on data variability[C]// Proceedings of the 1997 IEEE National Radar Conference. 1997: 263-268.
- [53] IERVOLINO P, GUIDA R. A novel ship detector based on the generalized-likelihood ratio test for SAR imagery[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 10(8): 3616-3630.
- [54] WANG R, XU F, PEI J, et al. An Improved Faster R-CNN Based on MSER Decision Criterion for SAR Image Ship Detection in Harbor[C]// IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2019). 2019: 1322-1325.
- [55] LIN Z, JI K, LENG X, et al. Squeeze and excitation rank faster R-CNN for ship detection in SAR images[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 16(5): 751-755.
- [56] KANG M, LENG X, LIN Z, et al. A modified faster R-CNN based on CFAR algorithm for SAR ship detection[C]// 2017 International Workshop on Remote Sensing with Intelligent Processing (RSIP). 2017: 1-4.

- [57] KANG M, JI K, LENG X, et al. Contextual region-based convolutional neural network with multilayer fusion for SAR ship detection[J]. Remote Sensing, 2017, 9(8): 860.
- [58] REN S, HE K, GIRSHICK R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[C]//Advances in neural information processing systems. 2015: 91-99.
- [59] JIAO J, ZHANG Y, SUN H, et al. A densely connected end-to-end neural network for multiscale and multiscene SAR ship detection[J]. IEEE Access, 2018, 6: 20881-20892.
- [60] ZHANG T, ZHANG X, SHI J, et al. Depthwise Separable Convolution Neural Network for High-Speed SAR Ship Detection[J]. Remote Sensing, 2019, 11(21): 2483.
- [61] CUI Z, LI Q, CAO Z, et al. Dense Attention Pyramid Networks for Multi-Scale Ship Detection in SAR Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(11): 8983-8997.
- [62] LI Q, MIN R, CUI Z, et al. Multiscale Ship Detection Based On Dense Attention Pyramid Network in Sar Images[C]//IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2019). 2019: 5-8.
- [63] LI Z, YOU Y, LIU F. Multi-Scale Ships Detection in High-Resolution Remote Sensing Image Via Saliency-Based Region Convolutional Neural Network[C]//IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2019). 2019: 246-249.
- [64] CHEN C, HU C, HE C, et al. SAR Ship Detection Under Complex Background Based on Attention Mechanism[C]//Chinese Conference on Image and Graphics Technologies. 2019: 565-578.
- [65] AN Q, PAN Z, YOU H. Ship detection in Gaofen-3 SAR images based on sea clutter distribution analysis and deep convolutional neural network[J]. Sensors, 2018, 18(2): 334.
- [66] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: Single shot multibox detector[C]//European Conference on Computer Vision (ECCV). 2016: 21-37.
- [67] WANG Y, WANG C, ZHANG H, et al. Combing Single Shot Multibox Detector with transfer learning for ship detection using Chinese Gaofen-3 images[C]//2017 Progress in Electromagnetics Research Symposium-Fall (PIERS-FALL). 2017: 712-716.

- [68] WANG Z, DU L, MAO J, et al. SAR target detection based on SSD with data augmentation and transfer learning[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 16(1): 150-154.
- [69] 杜兰, 刘彬, 王燕, 等. 基于卷积神经网络的 SAR 图像目标检测算法[J]. 电子与信息学报, 2016, 38(12): 3018-3025.
- [70] CHEN C, HE C, HU C, et al. A deep neural network based on an attention mechanism for SAR ship detection in multiscale and complex scenarios[J]. IEEE Access, 2019, 7: 104848-104863.
- [71] CHEN C, HE C, HU C, et al. MSARN: A Deep Neural Network Based on an Adaptive Recalibration Mechanism for Multiscale and Arbitrary-Oriented SAR Ship Detection[J]. IEEE Access, 2019, 7: 159262-159283.
- [72] LI Y, HUANG Q, PEI X, et al. RADet: Refine Feature Pyramid Network and Multi-Layer Attention Network for Arbitrary-Oriented Object Detection of Remote Sensing Images[J]. Remote Sensing, 2020, 12(3): 389.
- [73] HOU X, AO W, XU F. End-to-end Automatic Ship Detection and Recognition in High-Resolution Gaofen-3 Spaceborne SAR Images[C]//IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium. 2019: 9486-9489.
- [74] AO W, XU F, LI Y, et al. Detection and discrimination of ship targets in complex background from spaceborne ALOS-2 SAR images[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(2): 536-550.
- [75] XU F, WANG H, SONG Q, et al. Intelligent Ship Recongnition from Synthetic Aperture Radar Images[C]//IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2018). 2018: 4387-4390.
- [76] AO W, XU F, QIAN Y, et al. Feature clustering based discrimination of ship targets for SAR images[J]. The Journal of Engineering, 2019, 2019(20): 6920-6922.
- [77] ZHOU F, FAN W, SHENG Q, et al. Ship detection based on deep convolutional neural networks for polsar images[C]//IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium. 2018: 681-684.
- [78] CHEN S W, TAO C S, WANG X S, et al. Polarimetric SAR targets detection and classification with deep convolutional neural network[C]//2018 Progress in Electromagnetics Research Symposium (PIERS-Toyama). 2018: 2227-2234.

- [79] CLOUDE S R, POTTIER E. A review of target decomposition theorems in radar polarimetry[J]. IEEE Transactions on Geoscience and Remote Sensing, 1996, 34(2): 498-518.
- [80] KROGAGER E, BOERNER W M, MADSEN S N. Feature-motivated Sinclair matrix (sphere/diplane/helix) decomposition and its application to target sorting for land feature classification[C]//Wideband Interferometric Sensing and Imaging Polarimetry:vol. 3120. 1997: 144-154.
- [81] CAMERON W L, YOUSSEF N N, LEUNG L K. Simulated polarimetric signatures of primitive geometrical shapes[J]. IEEE Transactions on Geoscience and Remote Sensing, 1996, 34(3): 793-803.
- [82] TOUZI R, CHARBONNEAU F. Characterization of target symmetric scattering using polarimetric SARs[J]. IEEE Transactions on Geoscience and Remote Sensing, 2002, 40(11): 2507-2516.
- [83] HUYNEN J R. Phenomenological theory of radar targets[J]., 1970.
- [84] VAN ZYL J J. Unsupervised classification of scattering behavior using radar polarimetry data[J]. IEEE Transactions on Geoscience and Remote Sensing, 1989, 27(1): 36-45.
- [85] FREEMAN A, DURDEN S L. A three-component scattering model for polarimetric SAR data[J]. IEEE Transactions on Geoscience and Remote Sensing, 1998, 36(3): 963-973.
- [86] YAMAGUCHI Y, MORIYAMA T, ISHIDO M, et al. Four-component scattering model for polarimetric SAR image decomposition[J]. IEEE Transactions on Geoscience and Remote Sensing, 2005, 43(8): 1699-1706.
- [87] YAMAGUCHI Y, YAJIMA Y, YAMADA H. A four-component decomposition of POLSAR images based on the coherency matrix[J]. IEEE Geoscience and Remote Sensing Letters, 2006, 3(3): 292-296.
- [88] CLOUDE S R, POTTIER E. An entropy based classification scheme for land applications of polarimetric SAR[J]. IEEE Transactions on Geoscience and Remote Sensing, 1997, 35(1): 68-78.
- [89] HOLM W A, BARNES R M. On radar polarization mixed target state decomposition techniques[C]//Proceedings of the 1988 IEEE National Radar Conference. 1988: 249-254.

- [90] CHEN S W, OHKI M, SHIMADA M, et al. Deorientation effect investigation for model-based decomposition over oriented built-up areas[J]. IEEE Geoscience and Remote Sensing Letters, 2012, 10(2): 273-277.
- [91] SATO A, YAMAGUCHI Y, SINGH G, et al. Four-component scattering power decomposition with extended volume scattering model[J]. IEEE Geoscience and Remote Sensing Letters, 2011, 9(2): 166-170.
- [92] AN W, CUI Y, YANG J. Three-component model-based decomposition for polarimetric SAR data[J]. IEEE Transactions on Geoscience and Remote Sensing, 2010, 48(6): 2732-2739.
- [93] SHAN Z, ZHANG H, WANG C, et al. Four-component model-based decomposition of polarimetric SAR data for special ground objects[J]. IEEE Geoscience and Remote Sensing Letters, 2012, 9(5): 989-993.
- [94] SHAN Z, WANG C, ZHANG H, et al. Improved four-component model-based target decomposition for polarimetric SAR data[J]. IEEE Geoscience and Remote Sensing Letters, 2011, 9(1): 75-79.
- [95] 闫丽丽, 张继贤, 高井祥, 等. 一种适合方位建筑物的基于物理散射模型的极化 SAR 影像四分量分解方法[J]. 电子学报, 2015, 43(1): 203-208.
- [96] 杨波. 机载雷达多普勒波束锐化算法改进[J]. 现代雷达, (11): 53-55.
- [97] 邱天, 杜自成, 张涛. 提高雷达方位分辨力的多普勒波束锐化技术%Technique of Doppler Beam Sharpening for Enhancing Radar Azimuth Resolution[J]. 火控雷达技术, 034(2): 17-20.
- [98] 程剑熹, 周荫清. 多普勒波束锐化在群目标分辨中的应用[J]. 现代雷达, 1995(4): 9-14.
- [99] 翟庆林, 张军, 付强. 线性调频脉冲压缩技术及其在雷达系统中的应用[J]. 现代电子技术, 2007, 30(1): 17-19.
- [100] 韩晓东, 李晓明. 线性调频脉冲压缩雷达噪声干扰仿真研究[C]//江苏省电子学会 2010 年学术年会论文集. 2010.
- [101] 邢孟道, 保铮, 冯大政. 基于调幅-线性调频信号参数估计的机动目标成像方法[J]. 现代雷达, (6): 44-49.
- [102] 江魏好. DEM 辅助的星载 SAR 图像几何精校正技术[D]. 国防科学技术大学, 2016.

- [103] 陈吉庭. 星载 SAR 几何精校正技术[D]. 国防科学技术大学, 2015.
- [104] LEE J S. A simple speckle smoothing algorithm for synthetic aperture radar images[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1983(1): 85-89.
- [105] 高贵, 张军, 吕信明, 等. SAR 图像乘性噪声模型分析[J]. 信号处理, 2008, 24(2): 161-167.
- [106] UIJLINGS J R, VAN DE SANDE K E, GEVERS T, et al. Selective search for object recognition[J]. International journal of computer vision, 2013, 104(2): 154-171.
- [107] ZITNICK C L, DOLLÁR P. Edge boxes: Locating object proposals from edges[C]// European conference on computer vision. 2014: 391-405.
- [108] HE K, ZHANG X, REN S, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 37(9): 1904-16.
- [109] GIRSHICK R. Fast r-cnn[C]// Proceedings of the IEEE international conference on computer vision. 2015: 1440-1448.
- [110] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 779-788.
- [111] REDMON J, FARHADI A. YOLO9000: better, faster, stronger[C]// Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 7263-7271.
- [112] REDMON J, FARHADI A. YOLOv3: An Incremental Improvement[J]. [Online] Available: <https://arxiv.org/abs/1804.02767>,
- [113] LI J, QU C, SHAO J. Ship detection in SAR images based on an improved faster R-CNN[C]// 2017 SAR in Big Data Era: Models, Methods and Applications (BIGSAR-DATA). 2017: 1-6.
- [114] CAI Z, FAN Q, FERIS R, et al. A Unified Multi-scale Deep Convolutional Neural Network for Fast Object Detection[C]// ECCV. 2016.
- [115] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J].,

- [116] JIA Y, SHELHAMER E, DONAHUE J, et al. Caffe: Convolutional architecture for fast feature embedding[C]//Proceedings of the 22nd ACM international conference on Multimedia. 2014: 675-678.
- [117] PHILBIN J, CHUM O, ISARD M, et al. Object retrieval with large vocabularies and fast spatial matching[C]//IEEE conference on computer vision and pattern recognition. 2007: 1-8.
- [118] QIN X, ZHOU S, ZOU H, et al. A CFAR detection algorithm for generalized gamma distributed background in high-resolution SAR images[J]. IEEE Geoscience and Remote Sensing Letters, 2012, 10(4): 806-810.
- [119] HANSEN M G, JENSEN T K, LEHN-SCHIØLER T, et al. Empirical ship domain based on AIS data[J]. The Journal of Navigation, 2013, 66(6): 931-940.
- [120] CAI Z, VASCONCELOS N. Cascade r-cnn: High quality object detection and instance segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019.
- [121] 余映. 视觉注意计算模型设计及其应用研究[D]. 复旦大学, 2010.
- [122] 余映, 王斌, 张立明. 基于脉冲余弦变换的选择性视觉注意模型[J]. 模式识别与人工智能, 2010, 23(5): 616-623.
- [123] YU Y, WANG B, ZHANG L. Hebbian-based neural networks for bottom-up visual attention and its applications to ship detection in SAR images[J]. Neurocomputing, 2011, 74(11): 2008-2017.
- [124] YU Y, WANG B, ZHANG L. Hebbian-based neural networks for bottom-up visual attention systems[C]//International Conference on Neural Information Processing. 2009: 1-9.
- [125] COZZOLINO D, PARRILLI S, SCARPA G, et al. Fast adaptive nonlocal SAR despeckling[J]. IEEE Geoscience and Remote Sensing Letters, 2013, 11(2): 524-528.
- [126] LEE J S, KROGAGER E, AINSWORTH T L, et al. Polarimetric analysis of radar signature of a manmade structure[J]. IEEE Geoscience and Remote Sensing Letters, 2006, 3(4): 555-559.
- [127] SABRY R, VACHON P W. A unified framework for general compact and quad polarimetric SAR data and imagery analysis[J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 52(1): 582-602.

- [128] SONG Q, XU F, JIN Y. Reconstruction full-pol SAR data from single-pol SAR image using deep neural network[C]//2018 China International SAR Symposium (CISS). 2018: 1-4.
- [129] NAISHADHAM K, PIOUS J E. A robust state space model for the characterization of extended returns in radar target signatures[J]. IEEE Transactions on Antennas and Propagation, 2008, 56(6): 1742-1751.
- [130] 王雪松, 陈思伟. 合成孔径雷达极化成像解译识别技术的进展与展望. 2020. DOI: 10.12000/JR19109.
- [131] CHEN S W, WANG X S, XIAO S P. Urban damage level mapping based on co-polarization coherence pattern using multitemporal polarimetric SAR data[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(8): 2657-2667.
- [132] ZHAO L, YANG J, LI P, et al. Characterization of the coherent scattering induced by ridging patterns in agriculture by the use of polarimetric SAR imagery[J]. International journal of remote sensing, 2017, 38(12): 3502-3518.
- [133] VOORMANSIK K, JAGDHUBER T, ZALITE K, et al. Observations of cutting practices in agricultural grasslands using polarimetric SAR[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 9(4): 1382-1396.
- [134] WANG H, MAGAGI R, GOITA K. Comparison of different polarimetric decompositions for soil moisture retrieval over vegetation covered agricultural area[J]. Remote Sensing of Environment, 2017, 199: 120-136.
- [135] RASMUSSEN C E, WILLIAMS C K I. Gaussian Processes for Machine Learning[M]. 2005.
- [136] SRINIVAS N, KRAUSE A, SEEGER M, et al. Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design[C]//Proceedings of the 27th International Conference on Machine Learning. 2010: 1015-1022.
- [137] HE H, SIU W C. Single image super-resolution using Gaussian process regression[C]//CVPR 2011. 2011: 449-456.
- [138] PILLONETTO G, NICOLAO G D. Kernel selection in linear system identification Part I: A Gaussian process perspective[C]//IEEE Conference on Decision and Control and European Control Conference. 2011: 4318-4325.

- [139] ABDESSALEM A B, DERVILIS N, WAGG D J, et al. Automatic Kernel Selection for Gaussian Processes Regression with Approximate Bayesian Computation and Sequential Monte Carlo[J]. *Frontiers in Built Environment*, 2017, 3.
- [140] CAMPS-VALLS G, VERRELST J, MUNOZ-MARI J, et al. A survey on Gaussian processes for earth-observation data analysis: A comprehensive investigation[J]. *IEEE Geoscience and Remote Sensing Magazine*, 2016, 4(2): 58-78.
- [141] BAZI Y, MELGANI F. Classification of hyperspectral remote sensing images using Gaussian processes[C]//*IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2008)*:vol. 2. 2008: II-1013.
- [142] BAZI Y, MELGANI F. Gaussian process approach to remote sensing image classification[J]. *IEEE transactions on geoscience and remote sensing*, 2009, 48(1): 186-197.
- [143] CHANG C C, LIN C J. LIBSVM: a library for support vector machines[J]. *ACM TIST*, 2011, 2(3): 27.
- [144] YANG M Y, LIAO W, ROSENHAHN B, et al. Hyperspectral image classification using Gaussian process models[C]//*2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. 2015: 1717-1720.
- [145] CHEN K, JIAN P, ZHOU Z, et al. Semantic annotation of high-resolution remote sensing images via Gaussian process multi-instance multilabel learning[J]. *IEEE Geoscience and Remote Sensing Letters*, 2013, 10(6): 1285-1289.
- [146] CHEN K, ZHOU Z, HUO C, et al. A semisupervised context-sensitive change detection technique via Gaussian process[J]. *IEEE Geoscience and Remote Sensing Letters*, 2012, 10(2): 236-240.
- [147] CHANDOLA V, VATSAVAI R R. A gaussian process based online change detection algorithm for monitoring periodic time series[C]//*Proceedings of the 2011 SIAM International Conference on Data Mining*. 2011: 95-106.
- [148] RUIZ P, MATEOS J, CAMPS-VALLS G, et al. Bayesian active remote sensing image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, 52(4): 2186-2196.
- [149] KALANTARI L, GADER P, GRAVES S, et al. One-class Gaussian process for possibilistic classification using imaging spectroscopy[J]. *IEEE Geoscience and Remote Sensing Letters*, 2016, 13(7): 967-971.

- [150] VERRELST J, ALONSO L, CAMPS-VALLS G, et al. Retrieval of vegetation biophysical parameters using Gaussian process techniques[J]. IEEE Transactions on Geoscience and Remote Sensing, 2011, 50(5): 1832-1843.
- [151] VERRELST J, MUÑOZ J, ALONSO L, et al. Machine learning regression algorithms for biophysical parameter retrieval: Opportunities for Sentinel-2 and-3[J]. Remote Sensing of Environment, 2012, 118: 127-139.
- [152] ROELOFSEN H D, KOOISTRA L, van BODEGOM P M, et al. Mapping a priori defined plant associations using remotely sensed vegetation characteristics[J]. Remote sensing of environment, 2014, 140: 639-651.
- [153] SONG X, JIANG X, GAO J, et al. Gaussian Process Graph-Based Discriminant Analysis for Hyperspectral Images Classification[J]. Remote Sensing, 2019, 11(19): 2288.
- [154] YAN S, XU D, ZHANG B, et al. Graph embedding and extensions: A general framework for dimensionality reduction[J]. IEEE transactions on pattern analysis and machine intelligence, 2006, 29(1): 40-51.
- [155] VERRELST J, ALONSO L, CAICEDO J P R, et al. Gaussian process retrieval of chlorophyll content from imaging spectroscopy data[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2012, 6(2): 867-874.
- [156] LÁZARO-GREDILLA M, TITSIAS M K, VERRELST J, et al. Retrieval of biophysical parameters with heteroscedastic Gaussian processes[J]. IEEE Geoscience and Remote Sensing Letters, 2013, 11(4): 838-842.
- [157] VERRELST J, RIVERA J P, MORENO J, et al. Gaussian processes uncertainty estimates in experimental Sentinel-2 LAI and leaf chlorophyll content retrieval[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2013, 86: 157-167.
- [158] CAMPS-VALLS G, SVENDSEN D, MARTINO L, et al. Vegetation Monitoring with Gaussian Processes and Latent Force Models[C]//EGU General Assembly Conference Abs.Vol. 19. 2017: 12478.
- [159] CAMPS-VALLS G, SVENDSEN D H, MARTINO L, et al. Physics-aware Gaussian processes for Earth observation[C]//Scandinavian Conference on Image Analysis. 2017: 205-217.

- [160] MORALES-ALVAREZ P, PÉREZ-SUAY A, MOLINA R, et al. Remote sensing image classification with large-scale Gaussian processes[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 56(2): 1103-1114.
- [161] SVENDSEN D H, MARTINO L, CAMPOS-TABERNER M, et al. Joint Gaussian processes for biophysical parameter retrieval[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 56(3): 1718-1727.
- [162] BLIX K, ESPESETH M M, ELTOFT T. Up-Scaling from Quad-Polarimetric to Dual-Polarimetric SAR Data Using Machine Learning Gaussian Process Regression[C]// IEEE International Co Geosci. Remote Sens. Symp. (IGARSS). Valencia, Spain, 2018: 7332-7335.
- [163] BLIX K, ESPESETH M M, ELTOFT T. Machine Learning simulations of quad-polarimetric features from dual-polarimetric measurements over sea ice[C]//EUSAR 2018: 12th European Conference on Synthetic Aperture Radar. Aachen, Germany, 2018: 1-5.
- [164] SCHULZ E, SPEEKENBRINK M, KRAUSE A. A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions[J]. Journal of Mathematical Psychology, 2018, 85: 1-16.
- [165] TORRES R, SNOEIJ P, DAVIDSON M, et al. The Sentinel-1 mission and its application capabilities[C]//Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Munich, Germany, 2012: 1703-1706.
- [166] ABDIKAN S, SANLI F B, USTUNER M, et al. Land cover mapping using Sentinel-1 SAR data[J]. The International Archive on Photogrammetry, Remote Sensing and Spatial Information Science, 2016, 41: 757.
- [167] BARRA A, MONSERRAT O, MAZZANTI P, et al. First insights on the potential of Sentinel-1 for landslides detection[J]. Geomatics, Natural Hazards and Risk, 2016, 7(6): 1874-1883.
- [168] XIAO L, ZHENG J W, ZHANG Z L, et al. Application of Google Earth Software in Remote Sensing Geology Teaching[J]. Chinese Geological Education, 2017, 52(2).
- [169] YAO W. Study Cases on Fast Compression Distance Based Data Visualization[C]// IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium. 2019: 5804-5807.

- [170] CILIBRASI R, VITÁNYI P M. Clustering by compression[J]. IEEE Transactions on Information theory, 2005, 51(4): 1523-1545.
- [171] CILIBRASI R L, et al. Statistical inference through data compression[J]., 2007.
- [172] CERRA D, DATCU M. A fast compression-based similarity measure with applications to content-based image retrieval[J]. Journal of Visual Communication and Image Representation, 2012, 23(2): 293-302.
- [173] AHONEN T, HADID A, PIETIKÄINEN M. Face recognition with local binary patterns[C]//Proceedings of European Conference and Computer Vision. Prague, Czech Republic, 2004: 469-481.
- [174] ARROSPIDE J, SALGADO L. Log-Gabor filters for image-based vehicle verification[J]. IEEE Transactions on Image Processing, 2013, 22(6): 2286-2295.
- [175] CUI S, DUMITRU C O, DATCU M. Ratio-detector-based feature extraction for very high resolution SAR image patch indexing[J]. IEEE Geoscience and Remote Sensing Letters, 2013, 10(5): 1175-1179.
- [176] CHEN J, SHAN S, HE C, et al. WLD: A robust local image descriptor[J]. IEEE Transactions on Pattern Analysis and Machine Intelligent, 2010, 32(9): 1705-1720.
- [177] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Vol. 1. Los Alamitos, Calif, 2005: 886-893.
- [178] HOTELLING H. Analysis of a complex of statistical variables into principal components.[J]. Journal on Education Psychology, 1933, 24(6): 417.
- [179] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, Hawaii, 2017: 4700-4708.
- [180] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, Nevada, 2016: 770-778.
- [181] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size[J].,

- [182] DONG M, LIU D, XIONG Z, et al. 3D CNN-Based Soma Segmentation from Brain Images at Single-Neuron Resolution[C]//IEEE International Conference on Image Processing (ICIP). Athens, Greece, 2018: 126-130.
- [183] DEMIR B, BRUZZONE L. Histogram-based attribute profiles for classification of very high resolution remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 54(4): 2096-2107.
- [184] DEMIR B, MINELLO L, BRUZZONE L. An effective strategy to reduce the labeling cost in the definition of training sets by active learning[J]. IEEE Geoscience and Remote Sensing Letters, 2013, 11(1): 79-83.
- [185] HU J, GHAMISI P, ZHU X. Feature extraction and selection of Sentinel-1 dual-pol data for global-scale local climate zone classification[J]. ISPRS International Journal of Geo-Information, 2018, 7(9): 379.

致 谢

时光荏苒，六年充实忙碌的博士求学生涯已成过往。回首来路，往事历历在目，个中甘苦，冷暖自知。毕业之际，仍惶恐于是否达到了导师的期望，但一路走来，孜孜以求，不敢懈怠，学业终成。曾经焦虑过，也迷茫过，感谢这段痛并快乐的时光，让我受益颇多。值此论文完成之际，谨向指导、关心和帮助过我的人们表示最诚挚的谢意！

首先，衷心地感谢我的导师熊惠霖教授的培育之恩。熊老师学识渊博，治学严谨，待人谦逊平和，对学生关爱有加，时常关心我的科研进展与课题研究状况，并为我提出具有建设性的指导意见。老师对我感兴趣的研究内容给予了很大的支持，每每聆听老师的教诲总能让我备受启发和鼓舞。也非常感谢熊老师多次不厌其烦地帮我修订学位论文，有幸成为熊老师的学生并聆听老师的教诲，是我博士生涯最宝贵的财富。

感谢郁文贤教授在博士期间对我的指导。郁老师志存高远，执著于行，具有很强的的人格魅力，对科研事业孜孜不倦、精益求精的工作精神深深地感染着我，更激励着我。郁老师常常勉励我们要做事有主见、敢啃硬骨头。这些殷切的话语始终鞭策着我不断前行，唯恐有负老师的教诲。

也非常感谢张增辉教授的悉心指导。张老师数学功底深厚，专业知识扎实，在学术上为我指点迷津，引导我深入课题研究，项目上为我提供了很多锻炼的机会，注重对学生综合能力的全方位培养，在生活上也对我关怀备至，让我万分感激。感谢张肇健教授倾授科技论文写作方法，并不辞劳苦，多次帮我修订论文，让我受益匪浅。感谢郭炜炜老师和柳彬老师将我引入雷达图像解译与机器学习的殿堂，为我答疑解惑，帮我确定博士课题研究方向。感谢徐昌庆老师、武元新老师、李东瀛老师、熊刚老师、项艳老师、龚玲老师、陈敏老师、马金玉老师、朱莉黎老师等对我的指导和建议，感谢国防科技大学周石琳教授、邹焕新教授、计科峰教授对我的帮助和关心，愿众师身体安康。感谢师门胡昊、蒋武扬、丁拥科、孙清扬、蒋林峰、张涛、陈新竹、刘泽宇、吴春晓、黄兰清、杜振、张福强、李博颖、朱仕恒、代梦、王伟、余文豪、祝胜男、陈慧元等的帮助，感谢我最知心的室友朱文婕博士的陪伴，感谢班级陈昕苑、董一琳、陈彦君等同学的陪伴，相处数载，欢笑实多，愿诸君平安顺遂。

感谢我的外籍导师——德国宇航中心遥感技术研究所的 Mihai Datcu 教授。Datcu 教授是遥感领域国际知名学者，他渊博的知识、严谨的治学态度，使我在科研道路上受益颇多。他总是乐于分享自己的研究成果，百忙之中时常不忘参与学生的课题讨论，对学生生事无巨细的耐心指导，让我深受鼓舞。感谢 Gottfried Schwarz 博士与我的思维碰撞，

身体欠佳的他不能常去研究所上班,但仍会抽空与我相约慕尼黑市的各大图书馆讨论课题进展,让我非常感动。感谢朱晓香教授在留学期间给予我学业上的指导与生活上的关心。感谢德国宇航中心微波技术研究所的 Ralf Horn 和 Rolf Scheiber 博士与我共同探讨 F-SAR 数据的解译研究与使用。感谢同课题组的姚伟博士在 SAR 图像理解上的探讨并帮助我完成 OpenSARUrban 数据可视化工作,感谢 Corneliu Octavian Dumitru 博士在学习上的探讨与帮助。特别感谢庞蕾老师与崔世勇博士一家在求学期间对我生活上的帮助。感谢国家留学基金委对我的资助,让我开阔了眼界,增长了学识。

感谢上海交通大学对我的培养,选择了交大,就选择了责任,我一定会牢记母校的教诲,用更加优异的成绩回报母校。感谢上海交通大学感知与导航研究所为我提供良好的科研环境、数据与硬件支持,同时也感谢师弟师妹们协助我完成 OpenSARUrban 测试库中光学部分的标注。感谢欧洲空间局提供 Sentinel 数据和 SNAP 软件工具,感谢德国宇航中心提供 TerraSAR-X 数据和 F-SAR 数据。感谢上海交通大学@weijianwen 等同学提供了交大学位论文 L^AT_EX 模板,为我在学位论文排版上节省了很多时间。

深深地感谢一直以来默默支持和关心我的父母和亲人,感谢您们无怨无悔的付出和一生不变的牵挂,以及给予我的莫大的包容与鼓励。感谢弟弟和弟妹对父母的陪伴与照顾,才能够使我异地求学更加安心。感谢男友一如既往的支持与鼓励。

最后,向百忙之中审阅本文而付出辛勤劳动的各位专家、学者表示诚挚的谢意!

攻读学位期间发表的学术论文

- [1] **ZHAO, J.**, DATCU, M., ZHANG, Z., XIONG, H., YU, W. Contrastive-Regulated CNN in the Complex Domain: A Method to Learn Physical Scattering Signatures From Flexible PolSAR Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(12):10116-10135. (SCI 检索: 000505701800046)
- [2] **ZHAO, J.**, ZHANG, Z., YAO, W., DATCU, M., XIONG, H., YU, W.. OpenSARUrban: A Sentinel-1 SAR Image Dataset for Urban Interpretation[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2020, 1(13): 187-203. (SCI 检索: 000526639900016)
- [3] **ZHAO, J.**, GUO, W., ZHANG, Z., YU, W.. A Coupled Convolutional Neural Network for Small and Densely Clustered Ship Detection in SAR Images. Science China Information Sciences, 2019, 62(4):42301. (SCI 检索: 000445217500001)
- [4] **ZHAO, J.**, ZHANG, Z., YU, W., TRUONG, T. K.. A Cascade Coupled Convolutional Neural Network Guided Visual Attention Method for Ship Detection from SAR Images. IEEE Access, 2018, 6:50693-50708. (SCI 检索: 000447021300001)
- [5] **ZHAO, J.**, DATCU, M., ZHANG, Z., XIONG, H., YU, W. Translation of Physical Scattering Signatures Across Radar Sensors[J]. IEEE Transactions on Geoscience and Remote Sensing, **Under Review**. (SCI 源, 在审)
- [6] **ZHAO, J.**, DATCU, M., ZHANG, Z., XIONG, H., YU, W. Learning Physical Scattering Patterns from PolSAR Images by Using Complex-valued CNN[C]. 2019 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2019), 2019, 10019-10022. Yokohama, Japan.
- [7] **ZHAO, J.**, DATCU, M., ZHANG, Z., XIONG, H., YU, W. Physical Radar Signature Extraction from Flexible PolSAR Images[C]. TerraSAR-X / TanDEM-X Science Team Meeting 2019. DLR Oberpfaffenhofen, Germany.
- [8] **ZHAO, J.**, GUO, W., LIU, B., ZHANG, Z., YU, W., CUI, S.. Preliminary Exploration of SAR Image Land Cover Classification with Noisy Labels[C]. 2017 IEEE International

- Geoscience and Remote Sensing Symposium (IGARSS 2017), 2017, 3274-3277. Fort Worth, TX, USA.
- [9] **ZHAO, J.**, GUO, W., CUI, S., ZHANG, Z., YU, W.. Convolutional Neural Network for SAR Image Classification at Patch Level[C]. 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2016), 2016, 945-948. Beijing, China.
- [10] DENG, Z., SUN, H., ZHOU, S., **ZHAO, J.**, LEI, L., ZOU, H.. Multi-Scale Object Detection in Remote Sensing Imagery with Convolutional Neural Networks[J]. ISPRS journal of photogrammetry and remote sensing, 2018, 145:3-22.
- [11] DENG, Z., SUN, H., ZHOU, S., **ZHAO, J.**. Learning Deep Ship Detector in SAR Images from Scratch[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(6):4021-4039.
- [12] DENG, Z., SUN, H., ZHOU, S., **ZHAO, J.**, ZOU, H.. Toward Fast and Accurate Vehicle Detection in Aerial Images Using Coupled Region-based Convolutional Neural Networks[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 10(8):3652-3664.
- [13] MUSTAFA, N., **ZHAO, J.**, LIU, Z., ZHANG, Z., YU, W.. Iron Ore Region Segmentation Using High-Resolution Remote Sensing Images Based on Res-U-Net[C]. 2020 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2020), 2020, **Accepted**.
- [14] GUO, W., ZHANG, Z., **ZHAO, J.**, YU, W.. Fast Topology Preserving PolSAR Image Superpixel Segmentation[C]. 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2016), 2016, 6894-6897. Beijing, China.
- [15] DING, Y., GUO, W., **ZHAO, J.**, LI, Y., XIANG, W., ZHANG, Z., YU, W.. SAR Image Classification Based on CRFs with Object Structure Priors[C]. 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2016), 2016, 971-974. Beijing, China.
- [16] ZHU, P., WEN, L., et al. VisDrone-DET2018: The Vision Meets Drone Object Detection in Image Challenge Results[C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018, 437-468. Munich, Germany.

- [17] WEN, L., ZHU, P., et al. Visdrone-sot2018: The Vision Meets Drone Single-object Tracking Challenge Results[C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018, 469-495. Munich, Germany.

攻读学位期间参与的项目

- [1] 国家自然科学基金重点项目：高分辨率目标认知理论与关键技术（No. 61331015, 2014-2018）
- [2] “十三五”预研项目：基于脑机协同的 XX 目标检测技术（No.18GFA-ZZ01-047, 2018-2020）

攻读学位期间参加的学科竞赛及获奖情况

- [1] 2017 年 CCCV 遥感图像目标提取技术挑战赛（中国计算机视觉大会组委会主办）：三等奖
- [2] 2017 年“眼神杯”遥感影像稀疏表征与智能处理算法大赛（国家自然科学基金委信息学部、空间信息网络基础理论与关键技术重大研究计划指导专家组主办）：优秀奖
- [3] 2017 年第一届“航天星图杯”全国高分软件大赛（高分辨率对地观测学术年会组委会主办）：
- 高分辨率可见光图像切片典型地物目标分类：一等奖
 - 高分辨率 SAR 图像切片典型地物目标分类：二等奖
 - 高分辨率可见光图像中建筑物目标自动提取和重建：三等奖
 - 高分辨率可见光图像中飞机目标自动识别：三等奖
 - 高分辨率可见光图像中车辆目标自动识别：三等奖
 - 高分辨率 SAR 图像中舰船目标自动识别：优秀奖
 - 高分辨率 SAR 图像中建筑物目标自动提取和重建：优秀奖
 - 图像语义标注：优秀奖
- [4] 2018 年第二届“航天星图杯”全国高分软件大赛（高分辨率对地观测学术年会组委会主办）：
- 高分辨率可见光图像中飞机目标自动识别：一等奖
 - 高分辨率可见光图像中舰船目标自动识别：一等奖
 - 高分辨率 SAR 图像中建筑物目标自动提取：二等奖
 - 高分辨率可见光图像中污水处理厂自动识别：三等奖
 - 高分辨率可见光图像中建筑物目标自动提取：三等奖
 - 高分辨率 SAR 图像中舰船目标自动识别：优秀奖